PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [4]: C12N 15/00, 9/00, 9/54 C12N 9/96	**A1**	**(11) International Publication Number:** WO 88/ 07578 **(43) International Publication Date:** 6 October 1988 (06.10.88)

(21) International Application Number: PCT/US88/01078

(22) International Filing Date: 30 March 1988 (30.03.88)

(31) Priority Application Number: 034,085

(32) Priority Date: 2 April 1987 (02.04.87)

(33) Priority Country: US

(71) Applicant: GENENTECH, INC. [US/US]; 480 Point San Bruno Boulevard, San Francisco, CA 94080 (US).

(72) Inventors: WELLS, James, Allen ; 65 Otay Avenue, San Mateo, CA 94403 (US). CARTER, Paul, John ; 2074 18th Avenue, San Francisco, CA 94116 (US).

(74) Agent: PASSE, James, G.; Genencor, Inc., 180 Kimball Way, San Francisco, CA 94080 (US).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.

(54) Title: SUBSTRATE ASSISTED CATALYSIS

B. AMYLOLIQUEFACIENS SUBTILISN ————
PEPTIDE SUBSTRATE — · — · — · —
CATALYTIC TRIAD (Asp^{32}, His^{64}, Ser^{221}) — — — —

SUBSTRATE–SUBTILISIN
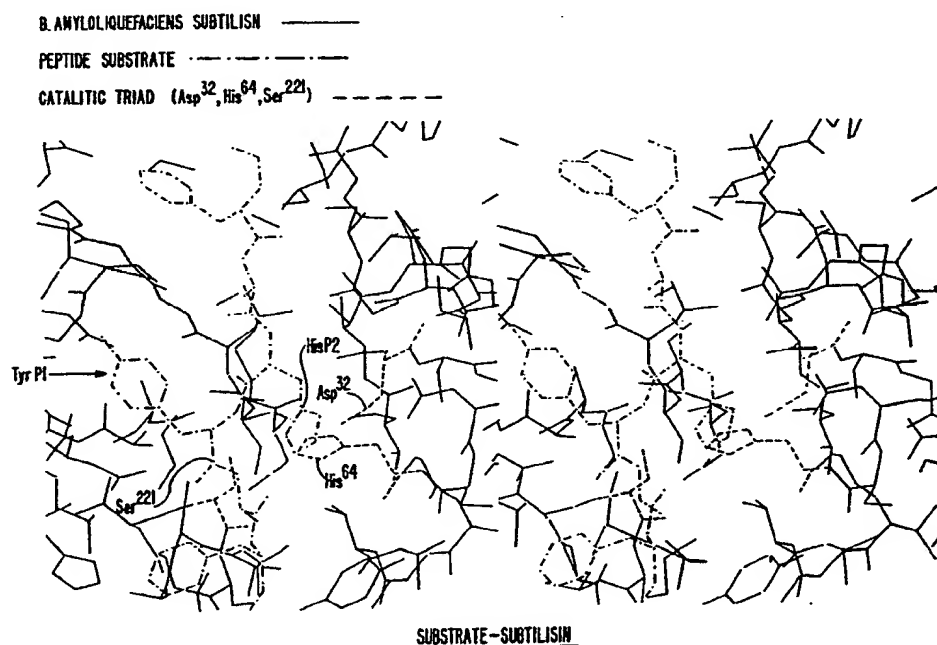
(57) Abstract

Novel enzyme mutants are disclosed which are derived from a precursor enzyme by replacing or modifying at least one catalytic functional group of an amino acid residue in a precursor enzyme. Such mutant enzymes have a catalytic preference for substrates which provide the replaced or modified catalytic group or its equivalent such that the substrate together with the enzyme mutant assists in its own catalysis.

SUBSTRATE ASSISTED CATALYSIS

This is a continuation-in-part of U.S. Patent Application Serial No. 846,627 filed April 1, 1986, which is a continuation-in-part of U.S Patent Application Serial No. 614,615 filed May 29, 1984 and

5   U.S. Patent Application Serial No. 858,594 filed April 30, 1986 which is a continuation-in-part of U.S. Patent Application Serial No. 614,612, 614,615, 614,617, 614,491 all filed May 29, 1984.

10   FIELD OF THE INVENTION
The present invention relates to novel enzyme mutants which are derived from a precursor enzyme by replacing or modifying at least one catalytic functional group of an amino acid residue in a precursor enzyme. Such

15   mutant enzymes have a catalytic preference for substrates which provide the replaced or modified functional group or its equivalent such that the substrate, in essence, together with the enzyme mutant, assists in its own catalysis.

20

PREFILING DISCLOSURES
Enzymes are polypeptides which catalyze a wide variety of chemical reactions. It is generally accepted that enzymatic catalysis requires that the substrate bind

25   to the enzyme in the region of the enzyme's active site such that the specific region being acted upon by the enzyme is distorted into a configuration

-2-

approximating the transition state of the reaction being catalyzed. In many cases the specific site of catalysis within the substrate must be oriented so that specific residues of the enzyme involved in catalysis can act on the bound and distorted substrate. Thus, within the active site, amino acid residues can generally be characterized as those primarily involved in substrate binding and hence determinative of substrate specificity and those involved primarily with the actual chemical catalysis, e.g., those involved in proton or electron transfer or nucleophilic or electrophilic attack on the substrate.

A wide variety of classical methods have been used to deduce the binding and catalytic residues in the active site of an enzyme. For example, the x-ray crystal structures of the serine endoprotease subtilisin containing covalently bound peptide inhibitors (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449), product complexes (Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303), and transition state analogs (Matthews, D.A., et al. (1975) J. Biol. Chem. 250, 7120-7126; Poulos, T.L. et al. (1976) J. Biol. Chem. 251, 1097-1103), which have been reported have provided information regarding the active site of subtilisin including the amino acid residues involved in substrate binding and catalytic activity. In addition, a large number of kinetic and chemical modification studies have been reported for subtilisin which have also aided in deducing the substrate binding and catalytic residues of subtilisin (Philip, M., et al. (1983) Mol. Cell. Biochem. 51, 5-32; Svendsen, I.B. (1976) Carlsberg Res. Comm. 41, 237-291; Markland, F.S. et al. (1971) In The Enzymes Ed. Boyer, P.D., Academic Press, New York, Vol. 3, pp. 561-608). In most cases where the chemical

-3-

modification was to a catalytic amino acid residue, the enzymatic activity of the enzymes modified was destroyed or severely impaired (Fersht, A. (1977) "Enzyme Mechanism and Structure", William Freeman,

5 San Francisco, California, pp. 201-205). In two reported examples, chemical modification of the active site serine of subtilisin resulted in the replacement of the serine-OH with -SH which produced a modified enzymatic activity (Neet, K.E., et al. (1968) J. Bio.

10 Chem. 248, 6392-6401; Polgar, L., et al. (1967) Biochemistry 6, 610-620). Most chemical modifications of catalytic residues, however, necessarily maintain or increase the effective side chain volume of the amino acid modified and consequently maintain or

15 decrease the effective volume within which the catalytic residues must function.

The recent development of various in vitro techniques to manipulate the DNA sequences encoding naturally-

20 occurring polypeptides as well as recent developments in the chemical synthesis of relatively short sequences of single and double stranded DNA has resulted in the reported synthesis of various enzymes wherein specific amino acid residues have been

25 substituted with different amino acids (Ulmer, K.M. (1983) Science 219, 666-671).

There are several reported examples where a catalytic residue of a particular enzyme has been substituted

30 with a different amino acid. Some of these references describe the replacement of a catalytic amino acid with an amino acid having a side chain functional group different from that of the catalytic amino acid being replaced e.g. substitution of a neutral polar

35 side chain moiety for a side chain moiety containing an acid group or substitution of one nucleophilic side

-4-

chain moiety with a different nucleophilic moiety. Others describe replacements where the side chain functional group of a catalytic residue remained constant but the position of that functional group was moved within the active site.

For example, Aspartate-102 of eucaryotic trypsinogen is reported to be a catalytic residue required for endoprotease activity. Roczniak, S.O., et al. (1985), J. Cell Biochem 9B (Abstracts) p. 87 briefly report the substitution of asparagine for aspartate at position 102. In this case, the carboxylate of aspartic acid was effectively substituted with the polar neutral side chain of asparagine which reportedly resulted in a dramatic decrease in kcat.

Dalbadie-McFarland, G., et al. (1982) PNAS (USA) 79, 6409-6413, report the inversion of the ser-thr diad of the β-lactamase gene contained in plasmid pBR322. This inversion resulted in the conversion of the catalytically active Serine-70 to Threonine and reportedly produced a mutant with an ampicillin-sensitive phenotype.

The substitution of Serine-70 in β-lacatamase with cysteine is reported by Sigal, I.S., et al. (1984) J. Bio. Chem. 259, 5327-5332. This replacement of an active site serine by a cysteine residue results in the net substitution of an -OH group by an -SH group, each of which can be effective nucleophiles. The thiol-containing β-lactamase reportedly catalyses the hydrolysis of β-lactams with a substrate specificity that is distinct from that of the wild type enzyme. For benzyl penicillin and ampicillin, the Km values are similar to wild type values although the kcat values are 1-2% that of a wild type enzyme. However,

-5-

when reacted with the cephalosporin nitrocefin, the Km
is greater than 10 fold that of the wild type and the
kcat is at least as large as the kcat for the wild
type enzyme.

5

In Strauss, et al. (1985) PNAS (USA) 82, 2272-2276,
triosphosphate isomerase was reportedly modified to
replace glutamic acid at position 165 with aspartic
acid.   This replacement does not alter the chemical
nature of the side chain at position 165 but rather
moves the catalytic carboxyl group at that position,
in essence, by the removal of a methylene group from
glutamic acid.   The kcat for different substrates was
dramatically altered by this mutation leading the
author to conclude that glutamic acid at position 165
is critical for proton shuttling during catalysis and
further suggesting that this residue makes only a
small contribution to the binding of the reaction
intermediates.

20

The substitution of Serine-102 in the active site of
alkaline phosphatase with cysteine is reported by
Ghosh, S.S., et al. (1986) Science 231, 154-148.   The
resulting thiol enzyme catalyzes the hydrolysis of a
variety of phosphate monoesters.   The authors
hypothesize, however, based on the observed catalytic
effeciency of the thiol containing enzyme, that the
serine to cysteine mutation results in a change in the
rate-determining step of catalysis from dephosphoryla-
tion to the formation of a phosphoryl-enzyme
intermediate.

The substitution of different amino acids for putative
catalytic residues in various enzymes has been
directed to the determination of whether these
residues are primarily involved in catalysis rather

-6-

than substrate binding. In several reported cases, however, the expected result was not obtained. In Gardell, S.J. et al. (1985) Nature 317, 551-555, Tyrosine-248 in carboxypeptidase A from rat was
5   substituted with phenylalanine. Tyrosine-248 had previously been thought to play a role in catalysis through its phenolic side chain. The particular substitution described removed the putative phenolic hydroxide moiety and substituted a phenyl moiety. The
10 authors report that the catalytic reactivity of the wild type enzyme compared to the substituted enzyme containing phenylalanine at position 248, for certain substrates, indicated that Tyrosine-248 was not obligatory for the hydrolysis of peptide substrates.
15 Rather, the authors suggests that the Tyrosine-248 hydroxyl group participates in substrate binding rather than catalysis.

Similarly, Threonine-113 in dihydrofolate reductase
20 from E.coli is a strictly conserved residue at the dihydrofolate binding site which interacts with a second conserved residue, Aspartate-27, via a hydrogen bond and presumably with the substrate dihydrofolate indirectly through a water molecule (Jin-Tann Chen, et
25 al. (1985) J. Cell. Biochem 29, 73-82). Since Aspartate-27 is also conserved and involved in catalysis, this suggested to the authors that Threonine-113 could be required for proton transfer during catalysis. The authors report the substitution
30 of Threonine-113 with valine and conclude that Threonine-113 is not involved in catalysis since there is no loss of catalytic efficiency upon substitution with valine.

35 Schultz, S.C., et al. (1986) PNAS (USA) 83, 1588-1592, report the substitution of threonine-71 in class A

β-lactamase with all possible amino acid substitutions to determine the role of this residue. Threonine-71 is a residue in the conserved triad Ser-Thr-Xaa-Lys. The results obtained by these authors suggests that Threonine-71 is not essential for binding or catalysis, as expected, but is important for stability of the β-lactamase protein.

Much of the work involving the substitution of different amino acids in various enzymes has been directed to the substitution of amino acid residues involved in substrate binding. Examples include the substitution of single amino acids within the active site of tyrosyl-tRNA synthetase (Cysteine-35→Serine, Winter, G. et al. (1982) Nature 299, 756-758; Cysteine-35→Glycine, Wilkinson, A.J. et al. (1983) Biochemistry 22, 3581-3586; and Threonine-51→Alanine and Threonine-51→Proline, Wilkinson A.J. et al. (1984) Nature 307, 187-188).

Other examples of substitutions of amino acids involved in substrate binding include a double mutant of tyrosyl-tRNA synthetase involving Cysteine-35→ Glycine together with Threonine-51→Proline (Carter, P.J. et al. (1984) Cell 38, 835-840); the substitution of glycine residues at positions 216 and 226 of rat pancreatic trypsin with alanine residues to produce two single substitutions and one double substitution (Craik, C.S. et al. (1985) Science 228, 291-297); and the substitution of various non-catalytic residues in dihydrofolate reductase (Villafranca, J.E., et al. (1983) Science 222, 782-788).

Paluh, J.L., et al. (1984) J. Biol. Chemistry 260, 1188-1894, report the substitution of Cysteine-84 with glycine in Serratia marcescens anthranilate synthase

-8-

Component II. They report that this replacement abolished the glutamine-dependent anthranilate synthase activity but not the ammonium-dependent activity of the enzyme. They also conclude that the mutation provides further evidence for the role of the active site Cysteine-84 in the glutamine amide transfer function of the enzyme. The authors also note, however, that the specific amino acid replacement might cause a relatively minor structural alteration that could abolish a glutamine binding or amide transfer independent of the function of Cysteine-84. It is not clear from this reference whether Cysteine-84 is a residue involved in binding, or actual catalysis.

The substitution of amino acid residue believed to be involved in transition state stabilization of various enzymes have also been reported. Such work has recently been summarized in Fersht, A.R., et al. (1986), Trends in Biochemical Sciences, 11, 321-325.

A reference in another field is Rossman, M.G., et al. (1985) Nature 317, 145-153 wherein the RNA of a human rhino virus is postulated to act as a proton acceptor for the autocatalytic cleavage of the viral coat protein VPO into VP2 and VP4.

The references discussed above are provided solely for their disclosure prior to the filing date of the present case, and nothing herein is to be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or priority based on earlier filed applications.

Based on the above references, however, it is apparent that those skilled in the art have focused on altering

enzyme specificity by changing binding residues. It
has heretofore not been recognized that residues
containing side chains directly involved in catalysis
can be substituted with residues containing smaller
5   and catalytically inactive side chains to produce
enzyme mutants which are catalytically active with
substrates which provide the catalytic function of the
replaced residue side chain. Thus, these enzyme
mutants have a substrate specificity which is
10   distinguished primarily at the level of catalysis
rather than substrate binding.

Accordingly, it is an object herein to provide enzyme
mutants wherein at least one catalytic group of an
15   amino acid residue of a precursor enzyme is replaced
or modified such that the thus formed mutant enzyme
has a preferred catalytic activity for a substrate
which is capable of providing the replaced or modified
catalytic function when in contact with the mutant
20   enzyme.

It is a further object to provide DNA sequences
encoding such enzyme mutants as well as expression
vectors containing such mutant DNA sequences.

25

Still further, another object of the present invention
is to provide host cells transformed with such vectors
as well as host cells which are capable of expressing
such enzyme mutants either intracellularly or
30   extracellularly.

A further object of the present invention is to
provide a catalytically active mutant enzyme substrate
complex wherein at least one of the catalytic
35   functional group, of the complex is provided by the
substrate.

-10-

Still further, an object of the present invention is to provide processes wherein the enzyme mutants of the invention are contacted used with modified substrates to bring about desired enzymatic catalysis.

SUMMARY OF THE INVENTION

The invention includes enzyme mutants not found in nature which are derived from a precursor enzyme by the replacement or modification of at least one catalytic group of an amino acid residue which when in contact with a selected region of a polypeptide substrate functions catalytically therewith. The enzyme mutant so formed is relatively inactive catalytically with the corresponding substrate as compared to the mutant's catalytical activity with a modified substrate formed by replacing or modifying a moiety in a selected region of the precursor enzyme's substrate. This selected region of the substrate is modified to include the catalytic group, or its equivalent which is replaced or modified in the precursor enzyme, such that the enzyme mutant is catalytically active with the modified substrate.

The invention also includes mutant DNA sequences encoding such mutant enzymes, expression vectors containing such mutant DNA sequences and host cells transformed with such vectors which are capable of expressing said enzyme mutants.

The invention also includes a catalytically active enzyme-substrate complex comprising an enzyme mutant and a modified substrate. The enzyme mutant is not found in nature and is derived from a precursor enzyme by the replacement or modification of at least one catalytic group of an amino acid residue which, when in contact with a selected region of a substrate for

-11-

the precursor enzyme, functions catalytically with
such substrate. The enzyme mutant so formed is
relatively inactive with the substrate for the
precursor enzyme as compared to the enzyme mutant's
5   catalytic activity with a modified substrate. The
modified substrate is formed by replacing or modifying
a moiety in the selected region of the precursor
enzyme's substrate. This selected region of the
substrate is modified to include the catalytic group,
10   or its equivalent, which is replaced or modified in
the precursor enzyme such that the enzyme mutant is
catalytically active with the modified substrate.

The invention further includes a process comprising
15   contacting an enzyme mutant and a modified substrate
to produce substrate assisted catalysis of the
modified substrate. In this aspect of the invention,
the enzyme mutant is the same as that defined for the
enzyme mutant-substrate complex of the invention.

20

BRIEF DESCRIPTION OF THE DRAWINGS
Fig. 1 is the DNA and amino acid sequence for B.
amylolquefaciens subtilisin.

25   Fig. 2 depicts catalytic residues of B.
amylolquefaciens subtilisin.

Figs. 3A and 3B depict the amino acid sequence of
subtilisin as obtained from various sources.

30

Fig. 3C depicts the conserved residues of B.
amylolquefaciens subtilisin when compared to other
subtilisin sequences.

35   Fig. 4 is a schematic diagram showing the substrate
binding cleft to subtilisin together with a substrate.

-12-

Fig. 5 is a stero view of B. amylolquefaciens subtilisin containing a modeled bound peptide substrate having the sequence L-Phe-L-Ala-L-His-L-Tyr-L-Gly-L-Phe representing residues P4 to P2' of the substrate.

Fig. 6 depicts the plasmid PS4.

Fig. 7A, 7B and 7C depict the pH dependence of hydrolysis of p-nitroanilide peptide substrates by Cys-24/Ala-64 subtilisin.

Fig. 8 depicts the hydrolysis of a polypeptide substrate by Cys-24/Ala-64 subtilisin.

Fig. 9 depicts a stero view of a complex between bovine trypsin and pancreatic trypsin inhibitor complex.

DETAILED DESCRIPTION OF THE INVENTION
The inventors have discovered that various catalytic groups in amino acid side chains in an enzyme can be replaced or modified to produce a mutant enzyme which is reactive with substrates which contain the replaced or modified catalytic group. The replaced or modified catalytic group is located in the substrate such that it is able to assist, with the mutant enzyme, in the catalysis of the modified substrate.

Specifically, B. amyloliquefaciens subtilisin, an alkaline bacterial protease, has been mutated by modifying the DNA encoding subtilisin to encode the substitution of the catalytic residue His-64 with alanine. As expected, kcat and the catalytic efficiency, as measured by kcat/Km, of this mutant enzyme was significantly reduced as compared to the

wild type subtilisin when contacted with substrates readily cleaved by wild type subtilisin. Surprisingly, various substrates containing histidine in the P2 position were preferred by the Ala-64 mutant subtilisin.

Previous studies have focused on altering enzyme specificity by changing residues that bind the substrate to the enzyme. The alternative approach described herein, termed "substrate-assisted catalysis", is applicable to a wide range of enzymes and substrates other than those specifically disclosed herein. In general, the invention is applicable to any enzyme in which part of the enzyme is removed and appropriately supplied by a similar functionality from a bound substrate. In this way substrates are distinguished primarily at the level of catalysis instead of binding, permitting the design of extremely specific enzyme mutants.

As used herein, "enzymes" are polypeptides which either alone or in conjunction with various co-factors catalyze a covalent change in a substrate. Enzymes can be categorized according to a systematic nomenclature and classification which has been adopted on the recommendation of the International Enzyme Commission. Thus, enzymes can be categorized as oxidoreductases (enzymes involved in oxidation reduction reactions), transferases (enzymes involved in the transfer of functional groups), hydrolases (enzymes involved in hydrolytic reactions), lyases (enzymes catalyzing addition reactions to double bonds), isomerases (enzymes involved in isomerization reactions) and ligases (enzymes involved in the formation of bonds with ATP cleavage). See, generally, Lehninger, A.L., Biochemistry, Worth

-14-

Publishers, Inc., New York, New York (1970), pp. 147-187.

A "precursor enzyme" refers to an enzyme in which a
5 catalytic amino acid residue can be replaced or
modified to produce a mutant enzyme. Typically, the
DNA sequence encoding the precursor enzyme may be
modified to produce a mutant DNA sequence which
encodes the substitution of one or more catalytic
10 amino acids in the precursor enzyme amino acid
sequence. Suitable modification methods are disclosed
herein and in EPO Publication No. 0130756 published
January 9, 1985. The precursor enzyme, however, can
also be modified by means other than recombinant DNA
15 technology to produce the mutant enzyme of the
invention.

A precursor enzyme may also be a recombinant enzyme
which refers to an enzyme for which its DNA has been
20 cloned or to an enzyme in which the cloned DNA
sequence encoding an enzyme is modified to produce a
recombinant DNA sequence which encodes the
substitution, deletion or insertion of one or more
amino acids in the sequence of a naturally occurring
25 enzyme. Suitable methods to produce such
modifications include those disclosed herein and in
EPO Publication No. 0130756. For example, the
subtilisin multiple mutant herein containing the
substitution of serine at amino acid residue 24 with
30 cysteine and the substitution of histidine at amino
and residue 64 with alanine can be considered to be
derived from the recombinant subtilisin containing the
substitution of cysteine for serine at residue 24.
The mutant thus is produced by the substitution of
35 alanine for histidine at residue 64 in the Cys-24
recombinant subtilisin.

-15-

Carbonyl hydrolases are enzymes which hydrolyze

$$\overset{O}{\overset{\|}{C}}$$

compounds containing C-X bonds in which X is oxygen or nitrogen. They include naturally-occurring carbonyl
5   hydrolases and recombinant or chemically synthesized carbonyl hydrolases. Naturally occurring carbonyl hydrolases principally include hydrolases, e.g. lipases and peptide hydrolases, e.g. subtilisins or metalloproteases. Peptide hydrolases include
10 α-aminoacylpeptide hydrolase, peptidylamino-acid hydrolase, acylamino hydrolase, serine carboxy-peptidase, metallocarboxypeptidase, thiol proteinase, carboxylproteinase and metalloproteinase. Serine, metallo, thiol and acid proteases are included, as
15 well as endo and exo-proteases.

Subtilisins are bacterial carbonyl hydrolases which generally act to cleave peptide bonds of proteins or peptides. As used herein, "subtilisin" means a
20 naturally occurring subtilisin or a recombinant subtilisin. A series of naturally occurring subtilisins is known to be produced and often secreted by various bacterial species. Amino acid sequences of the members of this series are not entirely
25 homologous. However, the subtilisins in this series exhibit the same or similar type of proteolytic activity. This class of serine proteases shares a common amino acid sequence defining a catalytic triad which distinguishes them from the chymotrypsin related
30 class of serine proteases. The subtilisins and chymotrypsin related serine proteases both have a catalytic triad comprising aspartate, histidine and serine. In the subtilisin related proteases the relative order of these amino acids, reading from the
35 amino to carboxy terminus is aspartate-histidine-serine. In the chymotrypsin related proteases the

-16-

relative order, however, is histidine-aspartate-serine. Thus, subtilisin herein refers to a serine protease having the catalytic triad of subtilisin related proteases.

5

Carbonyl hydrolases and their genes may be obtained from many procaryotic and eucaryotic organisms. Suitable examples of procaryotic organisms include gram negative organisms such as E. coli or pseudomonas and gram positive bacteria such as micrococcus or bacillus. Examples of eucaryotic organisms from which carbonyl hydrolase and their genes may be obtained include yeast such as S. cerevisiae, fungi such as Aspergillus sp., and mammalian sources such as, for example, Bovine sp. from which the gene encoding the carbonyl hydrolase chymosin can be obtained. As with subtilisins, a series of carbonyl hydrolases can be obtained from various related species which have amino acid sequences which are not entirely homologous between the members of that series but which nevertheless exhibit the same or similar type of biological activity. Thus, carbonyl hydrolase as used herein has a functional definition which refers to carbonyl hydrolases which are associated, directly or indirectly, with procaryotic and eucaryotic sources.

An "enzyme mutant" has an amino acid sequence which is derived from the amino acid sequence of a "precursor enzyme" and has a catalytic preference for a modified substrate as defined herein. The amino acid sequence of the enzyme mutant may be "derived" from the precursor amino acid sequence by the substitution of one or more catalytic amino acid residues of the precursor amino acid sequence. Suitable methods for such manipulation of the precursor DNA sequence include methods disclosed herein and in EPO

-17-

Publication No. 0130756. Other methods, for example, to directly modify the amino acid side chain of the precursor enzyme may be used provided they produce the catalytic preference for a modified substrate. A
5 "catalytic amino acid residue" is one which contains a catalytic group.

As used herein in connection with enzyme mutants, a "catalytic group" in an enzyme is a functional side
10 chain of an amino acid residue which undergoes a change in charge or chemical bonding state during a reaction sequence and which becomes regenerated at the end of the reaction sequence, or which interacts directly with such a functional side chain to
15 facilitate its change in charge or chemical bonding state. Catalytic groups typically participate in catalysis by interacting directly or indirectly as a nucleophile, electrophile, acid, base or electron transfer agent with the reactive site of a substrate.
20 Typical catalytic amino acid residues and their respective catalytic groups (shown in parentheses) include: Ser(-OH), Thr(-OH), Cys(-OH), Tyr(-OH), Lys($-NH_2$), Asp($-CO_2H$), Glu($-CO_2H$), His(imidazolyl) and Met($-SCH_3$). See Table I. Thus, for example,
25 catalytic groups for B. amyloliquefaciens subtilisin and as shown in Fig. 2 corresponding to the amino acid position numbers referred to in Fig. 1 comprise the side chains to the amino acids Asp-32, His-64 and Ser-221.
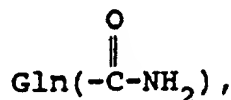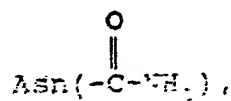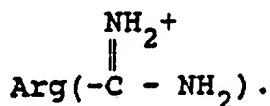
30

35

-18-

TABLE I

	Precursor Enzyme		Modified Substrate	
Catalytic Catalytic Residue	Preferred Amino Acid Residue Substitution	Alternate Amino Acid Residue Substitution	Catalytic Group	Equivalent Catalytic Group
His	Gly, Ala	Ser	Imidazoyl	$-NH_2$
Lys	Gly, Ala, Ser	Thr, Leu, Asn	$-NH_2$	Imidazolium
Ser	Gly	Ala	-OH	-SH
Thr	Gly	Ala	-OH	-SH
Cys	Gly	Ala	-SH	-HO
Asp	Gly, Ala	Ser	$-CO_2H$	Imadazoyl, Phenol
Glu	Gly, Ala Ser	Thr, Cys	$-CO_2H$	Imadazoyl, Phenol
Tyr	Gly, Ala Ser, Asn, Gln	Thr, Cys	Phenol	-OH, -SH Imidazolium
Met	Gly, Ala	Ser, Thr	$-S-CH_3$	-SH
Phe	Gly, Ala, Ser	Leu, Val	Phenyl	$-S-CH_3$, Phenol
Trp	Gly, Ala, Ser	Leu, Val	Indole	$-S-CH_3$, Phenol, Phenyl

-19-

As used herein in connection with the enzyme-substrate complexes or processes of the invention, a "catalytic group" in addition to the above definition, includes functional side chains of amino acid residues which
5   aid in stabilizing the transition state of a reaction by interacting directly or indirectly with a polarized or charged transition state. Such transition state stabilization is typically achieved by the formation of salt bridges or the creation of a dipole-dipole
10 interaction (e.g. hydrogen bond formation) between the transition   state   and   the   catalytic   residues stabilizing it. Typical catalytic amino acid residues involved in transistion state stabilization and their respective catalytic groups (shown in parenthesis)
15 include those catalytic residues of Table I:

$$\text{Asn}(-\overset{\overset{\textstyle O}{\|}}{C}-NH_2),$$

20

$$\text{Gln}(-\overset{\overset{\textstyle O}{\|}}{C}-NH_2),$$

and

25

$$\text{Arg}(-\overset{\overset{\textstyle NH_2^+}{\|}}{C}-NH_2).$$

(See   Table   II).   For   the   **B**.   _amyloliquefaciens_
30 subtilisin shown in Figs. 1 and 2, a catalytic residue involved in transition state stabilization is Asn-155 which   provides   a   hydrogen   bond   to   stabilize   the oxyanion of the tetrahydral intermediate shown in Fig. 2.

35

-20-

TABLE II

Precursor Enzyme			Modified Substrate	
Catalytic Catalytic Residue	Preferred Amino Acid Residue Substitution	Alternate Amino Acid Residue Substitution	Catalytic Group	Equivalent Catalytic Group
Asn	Gly, Ala,	Thr, Ser, Cys	$\overset{O}{\underset{\|}{-C}}-NH_2$	-OH Imidazoyl
Gln	Gly, Ala,	Thr, Ser, Cys	$\overset{O}{\underset{\|}{-C}}-NH_2$	-OH Imidazoyl
Arg	Gly, Ala,	Thr, Ser, Cys	$\overset{+NH_2}{\underset{\|}{-C}}-NH_2$	$-NH_3^+$ Imidazolium

-21-

Many enzymes are sufficiently characterized such that the catalytic groups of these enzymes (as defined above) are well known to those skilled in the art. However, for those enzymes which are not so characterized, the catalytic residues can be readily determined.

In this regard, amino acid replacement or chemical modification of catalytic groups (including those directly involved in catalysis and those involved in transition state stabilization) typically cause large disruptions in the catalytic step of the reaction (e.g. often measured by kcat) and little effect on the enzyme substrate dissociation constant (e.g. often measured by Km).

Thus, to determine whether a putative catalytic group is indeed catalytic, one skilled in the art can replace or modify the residue containing that group as described herein. If such substitution or modification abolishes or significantly reduces kcat, but does not substantially effect Km (e.g. increase/ decrease Km by a factor of 50 or preferably 10 or less), the side chain of the residue substituted is a catalytic group.

Structural methods such as x-ray crystallography can also be used to identify potential catalytic groups by their proximity to the site of the substrate chemical bond which becomes altered. Chemical, kinetic and nmr methods can also be useful in identifying catalytic groups by showing a change in their charge or chemical bonding properties during a reaction.

Alternatively, if a particular enzyme is not well characterized but is closely related to an enzyme

-22-

wherein one or more catalytic groups are already well-defined, the catalytic groups in that enzyme may be identified by determining its equivalent catalytic residues.

5

Thus, for example, a catalytic residue (amino acid) of a precursor carbonyl hydrolase is equivalent to a residue of B. amyloliquefaciens subtilisin if it is either homologous (i.e., corresponding in position in either primary or tertiary structure) or analagous to a specific residue or portion of that residue in B. amyloliquefaciens subtilisin (i.e., having the same or similar functional capacity to combine, react, or interact chemically).

15

In order to establish homology to primary structure in the above example, the amino acid sequence of a precursor carbonyl hydrolase is directly compared to the B. amyloliquefaciens subtilisin primary sequence and particularly to a set of residues known to be invariant in all subtilisins for which sequence is known (Figure 3C). After aligning the conserved residues, allowing for necessary insertions and deletions in order to maintain alignment (i.e., avoiding the elimination of conserved residues through arbitrary deletion and insertion), the residues equivalent to the catalytic amino acids (His-64, Asp-32, Ser-221, Asn-155) in the primary sequence of B. amyloliquefaciens subtilisin are defined. Alignment of conserved residues preferably should conserve 100% of such residues. However, alignment of greater than 75% or as little as 20% of conserved residues is also adequate to define equivalent residues.

35

-23-

For example, in Figure 3A the amino acid sequence of subtilisin from B. amyloliquefaciens B. subtilisin var. I168 and B. lichenformis (carlsbergensis) are aligned to provide the maximum amount of homology
5   between amino acid sequences. A comparison of these sequences shows that there are a number of conserved residues contained in each sequence. These residues are identified in Fig. 3C.

10 These conserved residues thus may be used to define the corresponding equivalent catalytic amino acid residues of B. amyloliquefaciens subtilisin in other carbonyl hydrolases such as thermitase derived from Thermoactinomyces. These two particular sequences are
15 aligned in Fig. 3B to produce the maximum homology of conserved residues. As can be seen there are a number of insertions and deletions in the thermitase sequence as compared to B. amyloliquefaciens subtilisin. Thus, the equivalent catalytic amino acid of Asn-155 in B.
20 amyloliquefaciens subtilisin in thermitase is the particular lysine shown beneath Asn-155.

Equivalent catalytic residues homologous at the level of tertiary structure for a precursor carbonyl
25 hydrolase whose tertiary structure has been determined by x-ray crystallography, are defined as those for which the atomic coordinates of 2 or more of the main chain atoms of a particular amino acid residue of the precursor carbonyl hydrolase and B. amyloliquefaciens
30 subtilisin (N on N, CA on CA, C on C, and O on O) are within 0.13nm and preferably 0.1nm after alignment. Alignment is achieved after the best model has been oriented and positioned to give the maximum overlap of atomic coordinates of non-hydrogen protein atoms of
35 the carbonyl hydrolase in question to the B. amyloliquefaciens subtilisin. The best model is the

-24-

crystallographic model giving the lowest R factor for experimental diffraction data at the highest resolution available.

$$\text{R factor} = \frac{\sum_{h}|\text{Fo}(h)| - |\text{Fc}(h)|}{\sum_{h}|\text{Fo}(h)|}$$

Equivalent catalytic residues which are functionally analogous to a catalytic residue of B. amyloliquefaciens subtilisin are defined as those amino acids of the precursor carbonyl hydrolases which may adopt a conformation such that they either alter, modify or contribute to catalysis in a manner defined and attributed to a specific catalytic residue of the B. amyloliquefaciens subtilisin as described herein. Further, they are those residues of the precursor carbonyl hydrolase (for which a tertiary structure has been obtained by x-ray crystallography), which occupy an analogous position to the extent that although the main chain atoms of the given residue may not satisfy the criteria of equivalence on the basis of occupying a homologous position, the atomic coordinates of at least two of the side chain atoms of the residue lie with 0.13nm of the corresponding side chain atoms of a catalytic group of B. amyloliquefaciens subtilisin. The three dimensional structures would be aligned as outlined above.

As used herein, a "substrate" refers to a substrate which is reactive with a precursor enzyme. For those enzymes which utilize polypeptides as substrate, the substrate is typically defined by an amino acid sequence which is recognized by the precursor enzyme

-25-

to bind the substrate therewith.   For example, a
substrate for trypsin contains the amino acid sequence

$$R$$
$$X - K - Y$$
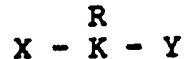
5 where X is any amino acid, Y is an amino acid except
Pro, R is Arg and K is Lys. Subtilisin is broadly
specific but will readily cleave a substrate with the
sequence FAAY↓AF (at the point designated by the
arrow) where F, A, Y are Phe, Ala and Tyr,
10 respectively, and Tyr occupies the P1 position (see
Fig. 4).   These residues, in general, are those
recognized by the enzyme to bind a particular
substrate and are referred to herein as a "selected
region" of the substrate. Of course, there may be a
15 wide range of polypeptides which are substrates for a
particular precursor enzyme. However, each of these
substrates will contain a selected region recognized
by the precursor enzyme.

20 In various aspects of the invention, the substrate may
be a non-proteinatious molecule such as a nucleic
acid, carbohydrate, a metabolite in a biological
pathway or an antibiotic.   In the case of nucleic
acids and carbohydrates the "substrate" can be
25 similarly defined as for a polypeptide substrate
except that these regions are defined not by amino
acid sequence but rather by nucleic acid sequence and
carbohydrate sequence, respectively.   Thus, for
example, restriction endonucleases recognize specific
30 DNA sequences and α amylase recognizes the α(1→4)
glycosidic linkage between glucose molecules in
amylose.

35 In the case of antibiotics, which usually are neither
polypeptides, nucleic acids or carbohydrates, there is
usually little problem in identifying the substrate.

-26-

Thus, for example, the substrates for β lactamases are penicillins and cephalosporans.

In general, substrates for a wide range of enzymes, including the selected region of such substrates, are known to or may be readily determined by those skilled in the art.

A "modified substrate" is a substrate wherein at least one moiety contained therein is replaced or modified to form a modified moiety which includes the catalytic group replaced or modified in a precursor enzyme or the equivalent of the catalytic group so replaced or modified. The catalytic group contained in the modified substrate is positioned such that upon binding with the mutant enzyme formed by modifying a particular precursor enzyme, the modified substrate provides a modified moiety which when in contact with the mutant enzyme provides the catalytic group replaced in the precursor enzyme or its equivalent. The thus formed enzyme mutant-modified substrate complex is thereby rendered catalytically active.

In the case of precursor enzymes having corresponding polypeptide substrates, the modified substrate is also a polypeptide which contains an amino acid sequence which binds to the mutant enzyme and which contains an amino acid residue within that selected region which has a side chain catalytic group which is the same or equivalent to the amino acid replaced or modified to form the mutant enzyme.

An "equivalent" catalytic group in a modified substrate is one which is capable of reacting, combining or interacting in the same or similar manner as that which was removed from or modified in the

-27-

precursor enzyme. Equivalent catalytic group refers
to a group having the ability to provide a similar or
equivalent catalytic role. It is not necessary that
an equivalent catalytic group provide equivalent
5   chemical structure. For example, if a catalytic His
residue is removed from the enzyme, an equivalent
catalytical group from the substrate would be an
imidazolyl group which may be donated (but not always
or exclusively) by a His side chain. If a catalytic
10 Ser residue is removed from the enzyme an equivalent
catalytic group from the substrate may be a hydroxyl
group which may be donated by Threonine or Tyrosine
side chain. In some cases the equivalent catalytic
group may not be identical to the original enzyme
15 group. Thus, an equivalent catalytic group for the
Serine-OH may be the Cysteine-SH. See Tables I and
II.

Upon binding with the enzyme mutant, the same or
20 equivalent catalytic group in the modified substrate
is capable of being positioned such that it is close
to the original position of the side chain of the
amino acid residue substituted or modified in the
precursor enzyme. In this manner the catalytic
25 function of the precursor enzyme can be reestablished
when the enzyme mutant binds a modified substrate.

The positioning of the catalytic group or equivalent
catalytic group within the selected region of a
30 substrate to form a modified substrate may be achieved
by substituting each of the amino acid residues within
the selected region with a different amino acid to
incorporate the catalytic or equivalent group in the
modified substrate at various positions. Such
35 modified substrates can be readily made by the methods
disclosed herein and by methods known to those skilled

-28-

in the art. Thereafter, these modified substrates are contacted with the particular mutant enzyme to determine which, if any, of the modified substrates are reactive with the enzyme mutant.

5

Alternatively, if the crystal structure of a particular enzyme or enzyme substrate complex is known, model building may be utilized to determine how a modified substrate should be constructed. Such model building may use, for example, a FRODO program (Jones, T.A. (1978) J. Appl. Crystallogr. 11, 268) in conjunction with an Evans and Sutherland PS300 graphics system. For example, the inventors have used such a program and graphic system to construct the stereo view of B. amylolquefaciens subtilisin shown in Fig. 5 containing a model bound peptide substrate having the sequence L-Phe-L-Ala-L-His-L-Tyr-L-Gly-L-Phe representing residues P4-P2' of the substrate.

A two-dimensional representation of the relationship between the subsites involved in subtilisin (S4 through S3') and the substrate residues involved in a substrate binding (P4 through P3') is shown in Fig. 4. Normally, the P2 position in the substrates typically reactive with subtilisin do not require histidine.

The model in Fig. 5 is based upon a 2.0 Å X-ray crystallographic study of product complexes bound to subtilisin. See e.g. Robertus, J.D. et al. (1972) Biochemistry 11, 4293; Poulos, T.L. et al. (1976) J. Biol. Chem. 251, 1097. The catalytic triad (Asp-32, His-64, and Ser-221) is shown with the His P2 side chain from the substrate superimposed upon the catalytic His-64. The distances between the OG of Ser-221 and the corresponding NE2 nitrogens from His-64 and the modeled P2 His side chain are 3.17 Å

and 3.17 Å, respectively. The distances between the OD2 of Asp-32 and the corresponding ND1 nitrogens from His-64 and the modeled P2 His side chain are 2.72 Å and 2.72 Å, respectively. The modeled distances

5   between the NE2 and ND1 nitrogens of the histidines are 1.39 Å and 1.35 Å, respectively. The hydrogen bond distances and dihedral angels for the stereo view of the complex of Fig. 4 are given in Table III as subtilisin model S1.

10

Likewise, the inventors have generated a stereo view of a complex between bovine trypsin and pancreatic trypsin inhibitor (PTI) complex in which the equivalent P2 substrate side chain (Cys-14 in PTI) is

15   replaced by His and superimposed upon His-57 in trypsin. The coordinates for the trypsin/trypsin inhibitor complex were taken from the Brookhaven Protein Data Bank entry 2PTC deposited by R. Huber and J. Deisenhofer, 9/82. See also Deisenhofer, J.,

20   et al. (1975) _Acta. Crystallogr._, Sect. B., _31_, 238. The catalytic triad of trypsin (Ser-195, His-57, Asp-102) is shown and the carbonyl carbon of Lys-15 at the P1 position in PTI is labeled. The hydrogen bond distances and dihedral angles for this stereo view in

25   Fig. 9 are given as trypsin model T1 in Table 3.

TABLE III

Pertinent bond angles and distances modeled for

30   substrate-assisted catalysis by a His P2 side chain in subtilisin or trypsin as depicted in Figs. 5 and 9, respectively. Dihedral angles for the His side chains are defined by $\chi 1(N-C\alpha-C\beta-C\gamma)$ and $\chi 2(C\alpha-C\beta-C\gamma-C\delta)$. The hydrogen bond angles $(N\epsilon 2(His)-H\delta(ser)-O\gamma(Ser))$

35   were   calculated   from   the   measured $C\beta(Ser)-O\gamma(Ser)-N\epsilon 2(His)$ angle, the $N\epsilon 2(His)-O\gamma(Ser)$

bond distance and the known O_γ(Ser)-Hδ(Ser) distance (0.96 Å) and the C_β(Ser)-O_γ(Ser)-Hδ(Ser) bond angle (108.5°) (Weiner, S.J. et al. (1984) J. Am. Chem. Soc. 106, 765). H-bond distances were measured between the catalytic Ser(O_γ) and Asp (Oδ1 and Oδ2) to the Nϵ2 and Nδ1, respectively, from the enzyme His or the substrate His P2. The distances are given between the enzyme His and the modeled substrate His P2 Nϵ2 and Nδ1 nitrogens. Model 1 (shown in Fig. 5 and 9 for subtilisin and trypsin, respectively) has the His P2 side chain optimized for H-bond distances between the imidazoyl nitrogen, Nϵ2 and Nδ1, to the catalytic Ser and Asp, respectively. Model 2 (graphic view not shown) has idealized χ angles for the His P2 side chain.

	Angles			Distances (Å)				
	Dihedral χ1	χ2	H-bond (Ser-His)	Nϵ2(His) →O_γ(Ser)	Nδ1(His)→ Oδ1(Asp)	or Oδ2(Asp)	Catalytic Nϵ2/Nϵ2	His-His P2 Nδ1/Nδ1
Subtilisin Catalytic His[64] (actual)	-167°	85°	148°	3.17	3.34	2.72	-	-
His P2 side chain Model S1	-164°	-50°	149°	3.17	3.55	2.72	1.39	1.35
Model S2	-180°	-90°	144°	3.25	3.59	3.34	0.37	1.57
Trypsin Catalytic His[57] (actual)	71°	85°	170°	2.70	3.25	2.70	-	-
His P2 side chain Model T1	-155°	-79°	179°	2.78	4.78	3.28	0.98	2.10
Model T2	-180°	-90°	158°	2.48	5.09	3.76	0.58	2.09

In general, modified substrates may be naturally occurring substrates containing amino acid sequences which previously were not recognized by the precursor enzyme or other enzymes or may be recombinant

5   substrates. Thus, for example, in the former case the inventors have determined that the subtilisin mutant Cys-24/Ala-64 is reactive with the naturally occurring substrates inhibin (between residues 61 and 80) and ACTH (between residues 1 and 10).

10

In the latter case, the recombinant substrate is engineered to be reactive with a specific enzyme mutant. Such recombinant substrates include, for example, a fusion polypeptide containing a pro

15   sequence (such as the Trp LE sequence from E. coli) and a desired polypeptide. Such fusion polypeptides are typically generated by recombinant techniques to facilitate the expression and/or secretion of the recombinant polypeptide. However, in many instances,

20   the fused sequence is not cleaved from the desired polypeptide upon secretion or by other known methods (e.g. by relatively nonspecific chemical reactions, such as treatment with CNBr, hydroxylamine, etc.). This problem is overcome by the use of the enzyme

25   mutants of the present invention where the polypeptide sequence of such fusion polypeptides is modified at the juncture of the pro and polypeptide sequences to incorporate a cleavage site which is recognized by the mutant enzyme and which will assist in its own

30   catalysis to produce the desired polypeptide free of the pro sequence. Thus, the invention may be used to engineer a unique cleavage site recognized by a mutant enzyme at the juncture of the pro and polypeptide sequence which does not result in secondary cleavage

35   within the desired polypeptide.

-32-

In the case of enzyme mutants which do not act on polypeptide substrates, the modified substrate will consist of a substrate for a precursor enzyme which has been appropriately modified to contain a modified
5   moiety which is catalytic when in contact with the enzyme mutant. Such modified substrates can be designed by substrate modeling as described above using the three-dimensional x-ray crystal structure of a precursor enzyme or enzyme-substrate complex. The
10   construction of such modified substrates, of course, will depend upon the chemical nature of the modified substrate as determined by such modeling and could involve biochemical and/or chemical modification or synthesis of the modified substrate.

15
In determining how a catalytic group should be replaced or modified in a precursor enzyme, consideration must be given to the modified substrate with which the enzyme is targeted to be reacted with.
20   In general, since the modified substrate will be providing a catalytic group removed from the precursor enzyme, the amino acid residue in the precursor enzyme should be replaced or modified in such a way as to provide space for the modified moiety of the modified
25   substrate. Typically, this requires that the side chain of the precursor amino acid be reduced in volume so that the enzyme mutant can recieve the moiety of the modified substrate.

30   The mean amino acid volume of amino acids when contained within a protein and the mean side chain volume of such amino acids normalized to a zero side chain volume for glycine are shown in Table IV. As shown in Tables I and II, there are various preferred
35   and alternate amino acids which may be substituted for specific catalytic residues within the active site of

-33-

a precursor enzyme. In each case, the amino acid
being substituted for a catalytic residue has a mean
side chain volume which is smaller than the side chain
of the catalytic residue replace. In general, the
5   catalytic amino acid residue should be replaced with
an amino acid such that the mean side chain volume
change upon making the substitution is sufficient to
accommodate the catalytic group or equivalent
catalytic group of the modified substrate as
10   determined empirically or by modeling studies. Thus,
for example, the substitution of His-64 for Ala
increases the active site volume by approximately
75 $\overset{o}{A}^3$ (101 $\overset{o}{A}^3$.-26 $\overset{o}{A}^3$). This increase in volume,
however, is sufficient to accommodate the histidine at
15   residue P2 in a modified substrate which has a mean
side chain volume of 101 $\overset{o}{A}^3$.

20

25

30

35

-34-

TABLE IV

Amino Acid	Chothia[1] Mean Amino Acid Volume in Protein ($\mathring{A}^3$)	Mean Side Chain Volume[2] ($\mathring{A}^3$)
Gly	66	0
Ala	92	26
Ser	99	33
Cys	118	52
Pro	129	63
Thr	122	56
Asp	125	59
Val	142	76
Asn	135	69
Ile	169	103
Glu	155	89
Leu	168	102
Gln	161	95
His	167	101
Met	171	105
Phe	203	137
Lys	171	105
Tyr	207	141
Arg	202	136
Trp	238	172

[1] Chothia (1984) _Ann. Rev. Biochem._ _53_, 537

[2] Normalized to zero side chain volume for glycine.

In addition to providing sufficient space for the catalytic group or equivalent catalytic group of the modified substrate, the side chain functionality of the catalytic residue replaced in the precursor enzyme should be altered to facilitate the binding and

catalytic activity of the modified substrate. For example, where the side chain of a catalytic amino acid residue in the presursor enzyme contains positively or negatively charged polar groups these

5 amino acids should be replaced or modified to contain side chain which contain non-polar or uncharged polar groups. Such substitutions are summarized in Tables I and II.

10 "Expression vector" refers to a DNA construct containing a DNA sequence which is operably linked to a suitable control sequence capable of effecting the expression of said DNA in a suitable host. Such control sequences include a promoter to effect

15 transcription, an optional operator sequence to control such transcription, a sequence encoding suitable mRNA ribosome binding sites, and sequences which control termination of transcription and translation. The vector may be a plasmid, a phage

20 particle, or simply a potential genomic insert. Once transformed into a suitable host, the vector may replicate and function independently of the host genome, or may, in some instances, integrate into the genome itself. In the present specification,

25 "plasmid" and "vector" are sometimes used interchangeably as the plasmid is the most commonly used form of vector at present. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions

30 and which are, or become, known in the art.

The "host cells" used in the present invention generally are procaryotic or eucaryotic hosts which preferably have been manipulated by the methods

35 disclosed in EPO Publication No. 0130756 to render them incapable of secreting enzymatically active

endoprotease. A preferred host cell for expressing subtilisin is the Bacillus strain BG2036 which is deficient in enzymatically active neutral protease and alkaline protease (subtilisin). The construction of

5   strain BG2036 is described in detail in EPO Publicatin No. 0130756 and further described by Yang, M.Y., et al. (1984) J. Bacteriol. 160, 15-21. Such host cells are distinguishible from those disclosed in PCT Publication No. 03949 wherein enzymatically inactive

10 mutants of intracellular proteases in E. coli are disclosed. Other host cells for expressing subtilisin include Bacillus subtilis I168 (EPO Publication No. 0130756).

15 Host cells are transformed or transfected with vectors constructed using recombinant DNA techniques. Such transformed host cells are capable of either replicating vectors encoding the enzyme mutants or expressing the desired enzyme mutant. In the case of

20 vectors which encode a pre or prepro form of the enzyme mutant, such mutants, when expressed, are typically secreted from the host cell into the host cell medium.

25 "Operably linked" when describing the relationship between two DNA regions simply means that they are functionally related to each other. For example, a presequence is operably linked to a peptide if it functions as a signal sequence, participating in the

30 secretion of the mature form of the protein most probably involving cleavage of the signal sequence. A promoter is operably linked to a coding sequence if it controls the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence

35 if it is positioned so as to permit translation.

-37-

The genes encoding the naturally-occurring precursor enzyme may be obtained in accord with the general methods described in EPO Publication No. 0130756 or by other method known to those skilled in the art. As can be seen from the examples disclosed in EPO Publication No. 0130756, the methods generally comprise synthesizing labelled probes having putative sequences encoding regions of the enzyme of interest, preparing genomic libraries from organisims expressing the enzyme, and screening the libraries for the gene of interest by hybridization to the probes. Positively hybridizing clones are then mapped and sequenced.

The cloned enzyme is then used to transform a host cell in order to express the enzyme. The enzyme gene is then ligated into a high copy number plasmid. This plasmid replicates in hosts in the sense that it contains the well-known elements necessary for plasmid replication: a promoter operably linked to the gene in question (which may be supplied as the gene's own homologous promotor if it is recognized, i.e., transcribed, by the host), a transcription termination and polyadenylation region (necessary for stability of the mRNA transcribed by the host from the hydrolase gene in certain eucaryotic host cells) which is exogenous or is supplied by the endogenous terminator region of the hydrolase gene and, desirably, a selection gene such as an antibiotic resistance gene that enables continuous cultural maintenance of plasmid-infected host cells by growth in antibiotic-containing media. High copy number plasmids also contain an origin of replication for the host, thereby enabling large numbers of plasmids to be generated in the cytoplasm without chromosonal limitations. However, it is within the scope herein to integrate

-38-

multiple copies of the hydrolase gene into host genome. This is facilitated by procaryotic and eucaryotic organisms which are particularly susceptible to homologous recombination.

5

Once the precursor enzyme gene has been cloned, a number of modifications are undertaken to enhance the use of the gene beyond synthesis of the naturally-occurring precursor enzyme. Such modifications
10   include the production of recombinant precursor enzymes (as disclosed in EPO Publication No. 0130756) and the production of enzyme mutants.

The following cassette mutagenesis method may be used
15   to facilitate the construction and identification of the enzyme mutants of the present invention although other methods including site-directed mutagenesis may be used. First, the gene encoding the enzyme is obtained and sequenced in whole or in part. Then the
20   sequence is scanned for a point at which it is desired to make a mutation of one or more amino acids in the expressed enzyme. The sequences flanking this point are evaluated for the presence of restriction sites for replacing a short segment of the gene with an
25   oligonucleotide pool which when expressed will encode various mutants. Such restriction sites are preferably unique sites within the gene so as to facilitate the replacement of the gene segment. However, any convenient restriction site which is not
30   overly redundant in the gene may be used, provided the gene fragments generated by restriction digestion can be reassembled in proper sequence. If restriction sites are not present at locations within a convenient distance from the selected point (from 10 to 15
35   nucleotides), such sites are generated by substituting nucleotides in the gene in such a fashion that neither

-39-

the reading frame nor the amino acids encoded are changed in the final construction. The task of locating suitable flanking regions and evaluating the needed changes to arrive at two convenient restriction

5 site sequences is made routine by the redundancy of the genetic code, a restriction enzyme map of the gene and the large number of different restriction enzymes. Note that if a convenient flanking restriction site is available, the above method need be used only in

10 connection with the flanking region which does not contain a site.

Mutation of the gene in order to change its sequence to conform to the desired sequence is accomplished by

15 M13 primer extension in accord with generally known methods. Once the gene is cloned, the restriction sites flanking the sequence to be mutated are digested with the cognate restriction enzymes and a plurality of end termini-complementary oligonucleotide cassettes

20 are ligated into the gene. The mutagenesis is enormously simplified by this method because all of the oligonucleotides can be synthesized so as to have the same restriction sites, and no synthetic linkers are necessary to create the restriction sites.

25

In the disclosed embodiment, subtilisin was chosen as a model to test the concept of substrate-assisted catalysis. In the hydrolysis of peptide bonds by subtilisin, His-64 acts as a catalytic base in the

30 formation of an acyl-enzyme intermediate and as a catalytic acid in the subsequent deacyclation step. Stroud, R.M., et al. (1975), Proteases and Biological Control (Cold Spring Harbor Laboratory, New York), p. 13; Kraut, J. (1977) Ann. Rev. Biochem. 46, 331.

35

-40-

The catalytic triad of subtilisin is shown in Fig. 2. As can be seen, Ser-221, His-64 and Asp-32 are positioned to facilitate nucleophilic attack by the serine hydoxylate on the carbonyl of the scissile peptide bond. Several hydrogen bonds may also help to stabilize the transition state complex for the tetrahedral substrate intermediate. One hydrogen bond is between aspartate and the positively charged histidine, ND1. Kossiakoff, A.A., et al. (1981) Biochem. 20, 6462-6474. A second hydrogen bond forms between the scissile amide nitrogen of the substrate and the (NE2) proton on the histidine. A third set of hydrogen bonds forms between the enzyme and the oxyanion that is produced from the carbonyl oxygen of the substrate. This latter set of hydrogen bonds is formed differently by the mammalian serine proteases and substilisin. A fourth hydrogen bond appears to exist between the amide nitrogen of the peptide bond between P-1 and P-2 and the carbonyl oxygen of Ser-125. Specifically, x-ray crystallographic studies of chymotrypsin (Henderson, R. (1970) J. Mol. Biol. 54, 341) indicate that two hydrogen bonds form between the substrate oxyanion and two main-chain amide protons from the enzyme (Gly-193 and the catalytic Ser-195). Crystallographic studies of subtilisin (Robertus, et al. (1972) Biochem. 11, 4293-4303; Matthews, et al. (1975) J. Biol. Chem. 250, 7120-7126; Poulos, et al. (1976) J. Biol. Chem. 250, 1097-1103) show that two hydrogen bonds are also formed with the oxyanion; one hydrogen bond donor is from the catalytic Ser-221 main-chain amide while the other is from one of the NE2 protons of the Asd-155 side chain. See Fig. 2.

The model shown in Fig. 5 revealed that the delta and epsilon nitrogens of the histidine in position P2 in

the modified substrate can be superimposed within about an angstrom by the corresponding nitrogens of the catalytic His-64 (not shown in Fig. 4). This suggested that if the histidine in the catalytic triad
5 of subtilisin was replaced by an alanine using site directed mutagenisis, a histidine from the substrate may substitue for the missing catalytic group in the mutant enzyme.

10 Maturation of the primary subtilisin gene product (preprosubtilisin) to subtilisin in B. subtilis is believed to be mediated by autoproteolysis that involves trace amounts of active subtilisin (Power, S.D., et al. (1986) Proc. Natl. Acad. Sci USA 83,
15 3096). The His-64→Ala mutation caused a severe reduction in secretion of mature subtilisin. However, it was possible to process and subsequently purify the inactive Ala-64 mutant by co-culturing B. subtilis cells harboring the Ala-64 mutant gene with B.
20 subtilis cells carrying an active subtilisin gene ("helper").

Stringent precautions were taken to ensure the purification of Ala-64 subtilisin away from "helper"
25 subtilisin and any other contaminating proteases. Firstly, the mutant subtilisin was expressed in the B. subtilis host BG2036 described in EPO Publication No. 0130756, that was deficient in chromosomal copies of the genes for alkaline protease (subtilisin) and
30 neutral protease. Secondly, to minimize "helper" contamination the ratio of "helper" cells to Ala-64 cells in the fermentation culture was adjusted to 1:1,000. Thirdly, a functionally silent Ser-24→Cys mutation that is located on the surface of subtilisin
35 (Wells, J.A., et al. (1986) J. Bio. Chem. 261, 6564) was introduced into the Ala-64 mutant. This

-42-

accessible cysteine served as an affinity handle for purificiation of the Ala-64 mutant away from the non-cysteine containing "helper" on an activated thiol sepharose column. Finally, the active "helper" subtilisin contained a functionally silent Ala-48→Glu mutation that altered its electrophoretic mobility relative to Cys-24/Ala-64 on native and SDS polyacrylamide gels. After purification, the Cys-24/Ala-64 mutant was judged to be greater than 99% pure by silver stained SDS (Morrissey, J.H. (1981) Anal. Biochem. 11, 307; Laemmli, U.K. (1970) Nature 227, 680) and native polyacrylamide gel electrophoresis. These purification procedures, including the use of a helper subtilisin which is capable of electrophoretic separation from the subtilisin mutant, are not necessarily required to practice the present invention.

Example 1

Construction of helper subtilisin containing
a functionally silent Ala-48→Glu mutation.

The construction of pS4 is described in detail in EPO Publication No. 0130756. This plasmid is depicted in Fig. 6. pS4 contains 4.5 kb of sequence derived from pBS42 (solid line) and 4.4kb of sequence containing the B. amyloliquefaciens subtilisin gene and flanking sequences (dashed line). pBS42 was constructed as described in EPO Publication No. 0130756 and Band, L. and Henner, D.J. (1984) DNA 3, 17-21. It was digested with BamHI and ligated with Sau3A partially digested chromosomal DNA from B. amyloliquefaciens (ATCC No. 23844) as described in EPO Publication No. 0120756. pS4 was selected from this genomic library.

-43-

pS4-5, a derivative of pS4 made according to Wells, et al. (1983) Nucleic Acids Res. 11, 7911-7924, was digested with EcoRI and BamHI, and the 1.5 kb EcoRI-BamHI fragment recovered. This fragment was

5　ligated into replicative form M-13 mp9 which had been digested with EcoRI and BamHI (Sanger, et al., (1980) J. Mol. Biol. 143, 161-178; Messing, et al., (1981) Nucleic Acids Res. 9, 304-321; Messing, J. and Vieira, J. (1982) Gene 19, 269-276). The M-13 mp9 phage

10　ligations, designated M-13 mp9 SUBT, were used to transform E. coli strain JM101 (ATCC 33876) and single stranded phage DNA was prepared from a two mL overnight culture. An oligonucleotide primer was synthesized having the sequence

15

5'-GTAGCAGGCGGAGAATCCATGGTTCC-3

The primer included the sequence of the subtilisin gene fragment encoding amino acids 44 through 52

20　except that the codon normally encoding alanine was substituted with the codon GAA encoding glutamate; the serine codon at 49(AGC) was also converted to TCC to introduce a convenient NcoI site.

25　The primer (about 15 μM) was labelled with [^{32}P] by incubation with [γ^{32}p]-ATP (10 μL in 20 μL reaction) (Amersham 5000 Ci/mmol, 10218) and T$_4$ polynucleotide kinase (10 units) followed by non-radioactive ATP (100 μM) to allow complete phosphorylation of the

30　mutagenesis primer. The kinase was inactivated by heating the phosphorylation mixture to 68°C for 15 minutes.

The primer was hybridized to M-13 mp9 SUBT as modified

35　from Norris, et al., (1983) Nucleic Acids Res. 11, 5103-5112 by combining 5 μL of the labelled

-44-

mutagenesis primer (~3 μM), ~1 μg M-13 mp9 SUBT
template, 1 μL of 1 μM M-13 sequencing primer
(17-mer), and 2.5 μL of buffer (0.3 M Tris pH 8, 40 mM
$MgCl_2$, 12 mM EDTA, 10 mM DTT, 0.5 mg/ml BSA).   The
5   mixture was heated to 68°C for 10 minutes and cooled
10 minutes at room temperature.   To the annealing
mixture was added 3.6 μL of 0.25 mM dGTP, dCTP, dATP,
and dTTP, 1.25 μ of 10 mM ATP, 1 μL ligase (4 units)
and 1 μL Klenow (5 units). The primer extension and
10   ligation reaction (total volume 25 μl) proceeded 2
hours at 14°C. The Klenow and ligase were inactivated
by heating to 68°C for 20 minutes.   The heated
reaction mixture was digested with BamHl and EcoRI and
an aliquot of the digest was applied to a 6 percent
15   polyacrylamide gel and radioactive fragments were
visualized by autoradiography.   This showed the [^{32}P]
mutagenesis primer had indeed been incorporated into
the EcoRI-BamHl fragment containing the now mutated
subtilisin gene.

20

The remainder of the digested reaction mixture was
diluted to 200 μL with 10 mM Tris, pH 8, containing 1
mM EDTA, extracted once with a 1:1 (v:v)
phenol/chloroform mixture, then once with chloroform,
25   and the aqueous phase recovered.   15 μL of 5M ammonium
acetate (pH 8) was added along with two volumes of
ethanol to precipitate the DNA from the aqueous phase.
The DNA was pelleted by centrifugation for five
minutes in a microfuge and the supernatant was
30   discarded.   300 μL of 70 percent ethanol was added to
wash the DNA pellet, the wash was discarded and the
pellet lyophilized.

pBS42 was digested with BamHl and EcoRI and purified
35   on an acrylamide gel to recover the vector.   0.5 μg of
the digested vector, 0.1 μg of the above primer

-45-

mutated EcoRI-BamHI digested subtilisin genonic fragment, 50 μM ATP and 6 units ligase were dissolved in 20 μl of ligation buffer. The ligation went overnight at 14°C. The DNA was transformed into the
5 B. subtilis host BG2036.

Example 2

10 ## Construction of His-64->Ala Mutant Subtilisin

The B. amyloliquifaciens subtilisin gene on a 1.5kb EcoRI-BamHI fragment (Wells, J.A., et al., (1983) Nucleic Acids Res. 11, 7911-7925) was cloned into
15 M13mp11 (Messing, J. and Vieira, J., (1982) Gene 19, 269-276) to give M13mp11SUBT and single-stranded DNA isolated (Carter, P., et al., (1985) in "oligonucleotide site-directed mutagenesis in M13" Anglian Biotechnology Limited). The mutation
20 His64->Ala was constructed using the synthetic oligonucleotide HA64 (5' CAACAACTCCGCGGGAACTCAC 3') and the M13SUBT template using a previously described method (Carter, P., et al., (1985) Nucleic Acids Res. 13, 4431-4443). The astrisk in HA64 denote mismatched
25 to the wild-type sequence and underlined is a unique SACII restriction site.

The primer (HA64) was annealed to the single-stranded M13SUBT template extended for 12 hrs. at 4°C with DNA
30 polymerase I (Klenow fragment) in the presence of deoxynucleoside triphosphates and T4 DNA ligase (Carter, P., et al., (1985) Nucleic Acids Res. 13, 4431-4443). The M13 heteroduplex DNA was then transfected directly into the E. coli host BMH 71-18
35 mutL (Kramer, B., et al., (1984) Cell 38, 879-887). Mutant phage were identified by colony blot

-46-

hybridization screening as previously described (Carter, P.J., et al., (1984) Cell 38, 835-840).

Putative His64->Ala mutants were verified by dideoxy nucleotide sequencing (Sanger, F., et al., (1977) Proc. Natl. Acad. Sci. USA 77, 5463-5467) as modified by Bankier, A.T. and Barrell, B.G., (1983) in "Techniques in the life sciences" B5, Nucleic Acids Biochemistry, B508, 1, Elsevier, Ireland and designated M13mp11SUBT-Ala-64. The 1.5kb EcoRI-BamHI fragment from M13mp11SUBT-Ala-64 was isolated and ligated with the 3.7 kb EcoRI-BamHI fragment from the B. subtilis-E. coli shuttle vector pBS42 (Band, L. and Henner, D.J., (1984) DNA 3, 17-21). E. coli MM294 cells (Murray, N.E., et al., (1977) Mol. Gen. Genet. 150, 53) were transformed with the ligation mixture using a $CaCl_2$ procedure (Cohen, S.N., Chang, A.C.Y., and Hsu, L., (1972) Proc. Natl. Acad. Sci. USA 59, 2110-2114). Plasmid DNA was recovered from individual transformants using an alkaline-sodium dodecyl sulfate (SDS) procedure (Birboim, H.C. and Doly, J., (1979) Nucleic Acids Res. 7, 1513-1528 as modified by Burke, J.F. and Ish-Horowicz, D., (1982) Nucleic Acids Res. 10, 3821-3830) to generate pBS42SUBT-Ala-64. The Ala64 mutation was verified by restriction endonuclease digests of the plasmid DNA using the enzymes SacII and BamHI which generate a 0.9kb fragment.

Example 3

Construction of the double mutant Ser-24->Cys/His-64->Ala

The double mutant Ser-24->Cys-24/His-64->Ala was constructed from the single mutants pBS42SUBT-Cys-24

-47-

(Wells, J.A. and Powers, D.B., (1986) <u>J. Biol. Chem</u>. <u>261</u>, 6564-6570) and pBS42SUBT-Ala-64 (this document) by a 3-way ligation using the following fragments: 3.7kb EcoRI/BamHI from pBS42, 0.5kb EcoRI/ClaI from

5   pBS42SUBT-Cys-24 and the 1.0kb ClaI/BamHI from pBS42SUBT-Ala-64. The double mutant Cys-24/Ala-64 was identified by restriction endonuclease site markers introduced by the single mutations (His-64->Ala: new SacII site; Ser-24->Ala: Sau3A site removed) and

10   designated pBS42SUBT-Cys-24/Ala-64. The pBS42SUBT-Cys-24/Ala-64 plasmid was introduced into the <u>B. subtilis</u> host BG2036 (Anagostopoluos, C. and Spizizen, J., (1961) <u>J. Bacteriol</u>. <u>81</u>, 741-746) deficient in alkaline and neutral proteases (Yang, M.Y., <u>et al</u>.,

15   (1984) <u>J. Bacteriol</u>. <u>160</u>, 15-21).

<u>Example 4</u>

20   Co-culturing of Cys-24/Ala-64 and
          <u>Glu48 mutant subtilisins</u>

Mutant subtilisin genes were expressed in BG2036 by fermentation in shake flasks using 2 X TY media

25   (Miller, J.H., (1972) in "Experiments in Molecular Genetics", Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) containing 12.5 µg/ml chloroamphenicol at 37°C for 18-20 hrs. Co-cultures wre made by diluting Cys-24/Ala-64 cultures 1:100 and Glu-48 cultures

30   1:100,000 in 2 x TY containing 12.5 µg/ml chloramphenicol and grown at 37°C for 20-24 hrs. with vigorous aeration.

35

-48-

Example 5

Purification of Cys-24/Ala-64

5   Cultures (21) were centrifuged (8,000g, 15 min., 4°C)
and 3 volumes of ethanol (-20°C) added to the
supernatant. After centrifugation (8,000g, 15 min.,
4°C), the pellet was resuspended in 50mM Tris.HCl (pH
8.0), 5mM CaCl$_2$, 10mM dithiothrietol (DTT), 0.1mM
10 phenylmethylsulfonyl fluoride (PMSF). After
centrifugation (40,000g, 30 min., 4°C) the supernatant
was dialysed against 21 10mM 2-[N-morpholino]ethane-
sulfonic acid (MES) (pH 6.0), 5mM CaCl$_2$, 10mM DTT,
0.1mM PMSF (S buffer) overnight at 4°C. The dialysate
15 was passed over a 50ml DE52 (Whatman) column and
loaded on to 50ml CM Trisacryl (LKB) column.
Subtilisin was eluted with a 600ml gradient of S
buffer containing 0-100mM NaCl at 1.5 ml/min. Pooled
subtilisin containing fractions were dialysed against
20 21 deaerated 10mM MES (pH 6.0), 5mM CaCl$_2$, 100mM NaCl,
0.1mM PMSF (T buffer). Samples were loaded on to an
activated thiol sepharose matrix (Pharmacia) washed
extensively with T buffer and then eluted with T
buffer containing 20mM DTT. The eluate was
25 concentrated using Centricon 10 microconcentrators
(Amicon) and then transferred to 10mM MES (pH 6.0),
5mM CaCl$_2$, 10mM DTT, 0.1mM PMSF (U buffer) by gel
filtration using PD10 G25 (Pharmacia) columns. The
concentration of subtilisin was determined from the
30 measured absorbance at 280nm (E_{280} 0.1% = 1.17)
(Matsubara, H., et al., (1965) J. Biol. Chem. 240,
1125-1130). Aliquots of purified enzyme were flash
frozen in liquid nitrogen and then stored at -70 C.

35

-49-

Example 6

Preparative native gel electrophoresis

5   ¯1.5mg Cys-24/Ala-64 subtilisin in U buffer was
adjusted to 10mM phenyl boronate, 10% glycerol (v/v)
and 0.1% (w/v) methylene blue. The sample was
electrophoresed for ¯24 hrs. at 7W (constant) on a 10%
polyacrylamide gel (20cm x 20cm x 0.75cm) with
10   recirculating buffer. The running buffer and gel
contained 10mM phenyl boronate, 2mM $CaCl_2$, 5mM DTT 50
mM histidine and 60mM 3-[N-morpholino]propanesulfonic
acid (MOPS). The protein was diffusion-blotted on to
nitrocellulose (Hancock, K., and Tsang, V.C.W., (1983
15   Anal. Biochem. 133, 157-162 as modified by Carter, P.,
et al., (1986) Proc. Natl. Acad. Sci. USA 83,
1189-1192). Subtilisin was visualized after binding
rabbit anti-subtilisin antibody (Power, S.D., et al.,
(1986) Proc. Natl. Acad. Sci. USA 83, 3096-3100) then
20   horse raddish peroxidase conjugated protein A by using
the chromogenic substrate 3,3'-diaminobenzidine
tetrahydrochloride. Subtilisin containing gel slices
were placed in dialysis bags with 6ml running buffer
(omitting phenyl boronate) and electroeluted at 10mA
25   (constant) for 20 hrs. at 4°C. Recovered material was
concentrated and transferred to U buffer as for column
purified enzyme (above).

30   ## Example 7

Kinetic analysis of Cys24/Ala64

The kinetic parameters for Cys-24 and Cys-24/Ala-64
35   were determined against the substrates N-succinyl-L-
Phe-L-Ala-L-[X]-L-Phe-p-nitroanilide   (abbreviated

sFAXF-pNA), where X (P2 position) was Ala, Gln, or His
(Table I). The kinetic parameters for the Cys-24
enzyme are essentially identical to wild-type
subtilisin against these substrates indicating that
5 the Ser24→Cys mutation is kinetically silent. By
comparison, the His-64→Ala mutation causes a drop of
$\sim 10^6$ fold in kcat/Km against the Ala and Gln P2
substrates. Almost all of the decrease in catalytic
efficiency is caused by a decreased k_{cat} term (up to
10 10^6 times), although smaller but significant increases
appear in Km. Unlike wild-type or Cys-24 subtilisin,
the Cys-24/Ala-64 enzyme was completely resistant to
inhibition by the active site reagent,
phenylmethylsulfonyl-fluoride (PMSF). This suggests
15 the catalytic histidine is critical for stable
sulfonylation by PMSF. Although the proportion of
functional active sites in Cys-24/Ala-64 enzyme
preparations could not be determined directly by such
active site labeling, enzyme that was purified by
20 additional native gel electrophoresis (Example 6) had
identical kinetic parameters to Cys-24/Ala-64
described in Table V.

The data are consistent with His-64 being extremely
25 important in catalysis (persumably by proton transfer)
and only marginally important in substrate binding.
However, because we cannot be sure that acylation is
rate limiting for the Ala-64 mutant, as it is for the
wild-type enzyme (Wells, J.A. (1986) <u>Phil. Trans. R.</u>
30 <u>Soc. Lond. A</u>, <u>317</u>, 415-423), the relatively small
changes in Km may not reflect changes in the enzyme-
substrate dissociation constant (K_s) but rather a
shift in the rate determining step of the reaction
(Guttreund, et al. (1956) <u>Biochem J.</u>, <u>63</u>, 656). In
35 any case, the catalytic histidine contributes a factor

-51-

of about 10^6 to the total enzymatic rate enhancement (Table V).

The catalytic efficiency of Cys-24 toward the three P2
5   substrates are all within a factor of five of each
other. For the Cys-24/Ala-64 mutant, kcat/Km for the
Ala and Gln P2 substrates are essentially the same;
however, hydrolysis of the HisP2 substrate is 170 to
210 times more efficient, respectively. Essentially
10 all of the increase in kcat/Km for the His over the
Ala and Gln P2 substrates results from the kcat term
being larger by a factor of 2,000 and 500,
respectively. The larger Km values for the His and
Gln P2 substrates compared to Ala may reflect reduced
15 binding affinity resulting from a bulky group at P2.
Larger Km values are also observed for the Gln and His
substrates for the Cys-24 enzyme. Thus, the drop in
kcat/Km caused by the His-64→Ala mutation is partially
restored when cleaving a His P2 substrate. The net
20 effect is a marked increase in substrate preference
for a His P2 side-chain brought about at the level of
catalysis rather than binding. The non-enzymatic
hydrolysis rate of the HisP2 substrate is similar to
the Ala and Gln P2 substrates (Table V). Thus, the
25 His P2 substrate only becomes functionally active in
the context of the catalytic groups provided by the
enzyme.

The fact that the catalytic efficiency of the
30 Cys-24/Ala-64 mutant against the His P2 substrate is
5,000 fold below wild-type suggests the His from the
substrate P2 functions poorly in catalysis. This may
result from the His P2 making poor steric contacts
and/or improper alignment of the catalytic triad.
35 Indeed, the model of the His P2 side-chain does not
exactly match the catalytic His-64 in that the planes

-52-

of the histidines from the enzyme and substrate are almost perpendicular to each other (Fig. 5).

Example 8

pH Dependence of Peptide Bond
Hydrolysis by Cys^{24}/Ala^{64}

The pH dependence of k_{cat}/K_m for wild-type subtilisin shows a sigmoidal increase from pH 6 to 8 (Glazer, A.N. (1967), _J. Biochem._, 242 433) that reflects the titration of the catalytic His^{64} ($pK_a = 7.1 \pm 0.1$). The wild-type pH profile remains relatively flat over the range of 8-10 and declines thereafter (Ottesen, et al. (1970), _In Methods of Enzymology_ (Ed. Perleman, Acad. Press, N.Y., Vol 19, p. 199)).

Fig. 7 shows the pH dependence of hydrolysis of p-nitroanilide peptide substrates by Cys-24/Ala-64 subtilisin. Analysis of Cys-24/Ala-64 against sFAAF-pNA (Fig. 7A) was determined as in Table V except using 100 mM Tris.HCl or 100 mM 3-[cyclohexylamino]-1-propane sulfonic acid (CAPS) buffer. The data was fitted assuming a linear relationship with hydroxide ion concentration (solid lines in Figs. 7A and 7b). Analysis of Cys-24/Ala-64 with sFAHF-pNA (Fig. 7C) was determined as in Table V except using 100mM 3-[N-morpholino] propanesulfonic acid (MOPS) buffer (filled circles) or 100 mM Tris.HCl (open circles) and then normalizing the ionic strength using KCl. The data was fitted to a sigmoid relationship (solid line) using a least-squares fit procedure.

The pH dependence of kcat/Km is markedly different for the Cys-24/Ala-64 enzyme. For the sFAAF-pNA

substrate, there is an increase of 15 fold in the kcat/Km between pH 8 to 10 (Fig. 7A). The kcat/Km shows a linear dependence upon hydroxide ion concentration (Fig. 7B) suggesting that a hydroxide ion can act as a catalytic base in the absence of a catalytic histidine side chain. If one were to extrapolate from the increase in kcat/Km as a function of hydroxide concentration $(2 \times 10^4 \ s^{-1} M^{-2})$, to the kcat/Km for Cys-24 against this same Ala P2 substrate $(8 \times 10^5 \ s^{-1} M^{-1})$, then the equivalent concentration of the hydroxide ion would be about 40 M.

In contrast, the kcat/Km for hydrolysis of the sFAHF-pNA by Cys-24/Ala-64 shows a sigmoidal pH dependence between pH 6 and 8 (Fig. 2C) that is similar to wild-type subtilisin. The pK_a of the activity dependent group is 6.8 ± 0.1, and almost all of the pH dependent changes in kcat/Km result from changes in kcat (data not shown). For the sFAHF-pNA substrate, there is not a strong linear increase in kcat/Km with hydroxide above pH 8 as observed for hydrolysis of sFAAF-pNA. These data strongly suggest that the P2 histidine side-chain from the substrate can substitute functionally for the missing catalytic histidine 64.

The data presented in Table V (measured at pH 8.6) underestimate the substrate preference for His over Ala (and Gln) because the kcat/Km for the sFAHF-pNA is maximal at pH 8.0 (Fig. 7C), whereas for the sFAAF-pNA substrate it is significantly lower at pH 8.0 (Fig. 7B). Thus, for Cys-24/Ala-64 at pH 8.0, we estimate the substrate preference is up to ~400 times for the His P2 substrate over the corresponding Ala or Gln substrates.

-54-

TABLE V

Kinetic analysis of mutant subtilisin against the substrates, N-succinyl-L-Phe-L-Ala-L-X-L-Phe-ρ-nitroanilide, where X is Ala, Gln, or His. Six hydrolysis assays were performed simultaneously

5   against corresponding substrate blanks in 0.10 M Tris-HCl (pH 8.6), 10 mM DTT at 25±0.1°C using a Kontron unvikon 860 spectrophotometer. Initial reaction rates were determined from the increase in absorbance caused by the release of ρ-nitroaniline

10   group (ϵ_M^{410}=8,480 M^{-1}, cm^{-1} (DelMar, E.G., et al. (1979) Anal. Biochem. 99 316)) and fitted by linear regression to an Eadi-Hofstee plot to calculate V_{max} and Km kcat was calculated from V_{max}/[enzyme], using the spectrophotometrically determined enzyme

15   concentration (Matsubara, et al. (1965) J. Biol. Chem. 240, 1125). Enzyme concentrations in the assays were about 50 μg/mL for Cys-24/Ala-64 and 1 μg/mL for Cys-24. Standard errors in all determinations were below 20%. Slight variation in the absolute kinetic

20   values has been observed between batches of enzyme, but the relative values among substrates has remained constant.

Substrate	Cys-24			Cys-24/Ala-64			Non-enzymatic
P2 residue	kcat s^{-1}	Km μM	kcat/Km s^{-1}M^{-1}	kcat s^{-1}	Km μM	kcat/Km s^{-1}M^{-1}	hydrolysis rate s^{-}
Ala	8.1	10	8.0×10^5	8.1×10^{-6}	32	0.25	1.7×10^{-7}
Gln	7.0	39	1.8×10^5	3.0×10^{-5}	150	0.20	7.1×10^{-8}
His	4.6	23	2.0×10^5	1.6×10^{-2}	380	42	7.9×10^{-8}

-55-

Several lines of evidence indicate that the activity we attribute to the Cys-24/Ala-64 enzyme is not the result of any other protease contamination. Firstly, the extreme substrate preference for His at the P2

5   position is unlike wild-type subtilisin or any known Bacillus protease. Secondly, the mutant has Km values which are significantly different from wild-type subtilisin suggesting differences in the energetics of substrate binding and/or catalysis. Thirdly, the

10 mutant is completely resistant to inhibition by PMSF, unlike other serine proteases. In fact, the kinetic determinations for the Cys-24/Ala-64 mutant are routinely made in the presence of PMSF to exclude any possibility of active "helper" subtilisin (Table V,

15 Fig. 7). Fourthly, the substrate dependent pH profiles are unlike any protease we are aware of. Fifthly, preparations of Cys-24/Ala-64 are extremely pure from other contaminating proteins based by analysis on SDS and native gels (>99%). Finally, the

20 kinetic values determined for Cys-24/Ala-64 that was additionally purified by native gel electrophoresis (Example 6) are essentially the same as these reported in Table V.

25

Example 9

Hydrolysis of Polypeptide
Substrates by Cys-24/Ala-64

30
To further evaluate the specificity of the Cys-24/Ala-64 mutant in comparison with Cys-24, both enzymes were incubated with a 20 residue fragment of the inhibin B chain at pH 8.0. The choice of the

35 peptide was based upon the presence of two histidines (position 5 and 11) along with 16 different amino

-56-

acids, and a variety of large hydrophobic amino acids
that are preferred amino acids at the P1 position of
wild-type subtilisin (Estell, D.A., et al. (1986)
Science 233, 659). Fig. 8 shows the hydrolysis of the
5   inhibin peptide substrate TVINHYRMRGHSPFANKLSC by
Cys-24/Ala-64 subtilisin. This substrate (100 μg) was
digested with 10μg Cys-24/Ala-64 (Fig. 8A) or 0.13 μg
Cys-24 (Fig. 8B). Reaction mixtures were in a total
volume of 250 μL containing 20 mM Tris.HCl (pH 8.0),
10   10 mM dithiothreitol, 5% (V/V) dimethyl sulfoxide and
1 mM PMSF (Cys-24/Ala-64 only). After indicated times
at 37°C, digestion products (monitored at 214 nm) were
eluted from a reverse phase HPLC column (Waters, C18)
using a gradient (from left to right) of 0-50% (v/v)
15   acetonitrile in 0.1% (v/v) trifluoroacetic acid.

After a 2-hour incubation with Cys-24/Ala-64
(Fig. 8A), a ~120 fold molar excess of inhibin peptide
(peak a) was cleaved to greater than 95% completion
20   into two pieces (peaks b and c). Amino acid
composition analysis of these two peptide fragments
indicated cleavage had occurred between Tyr-6 and
Arg-7, as expected for substrate assisted catalysis by
His-5 located at the P2 position from cleavage site.
25   After ten fold longer digestion (20 hr.) a minor third
peak appeared (labelled X in Fig. 8A). Analysis
showed it to have the same composition as the
undigested inhibin peptide. This minor product also
appeared in a non-enzymatic blank incubation. No
30   digestion was observed at the second histidine site.

In contrast to the two fragments produced by
Cys-24/Ala-64, the Cys-24 enzyme produced at least
seven fragments (Fig. 8B) at a similar extent of
35   digestion of starting material (compare 5 min.
digestion with Cys-24 to 2 hr. digestion with

-57-

Cys-24/Ala-64). Although none of these seven
fragments were sequenced, the first two produced
eluted from the HPLC profile at the same positions as
peaks b and c in Fig. 8A. Digestion to 95% completion
5 of the starting peptide by Cys-24 (30 min. incubation,
Fig. 8C) produced more than ten different peptide
fragments.

Digestion experiments for this and five other peptides
10 are summarized in Table VI. A ten residue fragment of
human ACTH was quantitatively cleaved at a single site
by Cys-24/Ala-64. Amino acid composition analysis of
the two digetion products confirmed that the cleavage
had occurred with a His residue at the P2 position of
15 the substrate as expected. However digestion of this
peptide with Cys-24 also gave specific cleavage at
this position (not shown). This probably resulted
because Phe provides a very favorable P1 residue, and
the two short peptides liberated do not provide
20 effective substrates for subtilisin. The four other
peptides tested (three containing His and one which
did not) were not cleaved by Cys-24/Ala-64 but were
cleaved at several sites by the Cys-24 mutant (not
shown).

25

30

35

-58-

TABLE VI

Digestion of peptide substrates by Cys-24/Ala-64
subtilisin. Various synthetic peptides (200 µg) shown
were digested with Cys-24/Ala-64 (10 µg) in 20 mM
Tris-HCl (pH 8.0), 10 mM DTT, 5% (v/v) dimethyl
5 sulfoxide, 1 mM PMSF (250 µL total volulme) for 20 hr.
at 37°C. Digestion products were analyzed by reverse
phase HPLC as described in Fig. 3. Digestion products
recovered by HPLC were hydrolyzed for 24 hr. in 6 N
HCl, 1% (v/v) phenol before amino acid analysis using
10 nor-leucine as an internal standard. The Cys residues
in bovine insulin A and B chains were oxidized to
$CysSO_3H$. Sequences are designated by the single
letter amino acid code.

Peptide Source	Sequence	Cleavage Peptides with Cys-24/Ala-64
Inhibin β-chain Residues 61-80	TVINHYRMRGHSPFANLKSC	TVINHY + RMRGHSPFANLKSC
ACTH Residues 1-10	SYSMEHFRWG	SYSMEHF + RWG
Ubiquitin	CKESTLHLVLRLRGG	Not cleaved
Peptide C	GYEHFENLRRRAASFQGKY	Not cleaved
Bovine insulin B chain (oxidized)	FVNQHLCGSHLVEALYLVCGERGFFYTPKA	Not cleaved
Bovine insulin A chain (oxidized)	GIVEQCCASVCSLYQLENYCN	Not cleaved

-59-

These experiments establish much about the activity, specificity, and utility of the Cys-24/Ala-64 mutant. In addition to ρ-nitroanilide substrates, the enzyme is capable of cleaving normal peptide bonds. Unlike

5   the Cys-24 enzyme, the specificity of Cys-24/Ala-64 appears to be limited to sites containing a histidine side chain located at the P2 position of the cleavage site.   Furthermore,   additional   specificity determinants are required because not all His P2 sites

10   are cleaved. We believe that this reflects the normal specificity determinants in the wild-type enzyme. Peptide substrates were chosen to have His followed by a large hydrophobic amino acid which is preferred for the P1 site in subtilisin (Estell, D.A. et al. (1986)

15   Sci. 233, 659; Phillip, M. et al. (1983) Mol. Cell Biochem. 51, 5; Svendsen, I. (1976) Carlsberg Res. Commun. 41, 237).   Little is known about P1' specificity but the absence of cleavage at the other His sites may reflect the presence of a negatively

20   charged (Glu or Cys SO_3H), a β-branched (Val) or a proline all of which are very poorly hydrolyzed P1 amino acids (Estell, D.A. et al. (1986) Sci. 233, 659).   All literature references are expressly incorporated herein by reference.

25

Having described the preferred embodiments of the present invention, it will appear to those ordinarily skilled in the art that various modifications may be made to the disclosed embodiments and that such

30   modifications are intended to be within the scope of the present invention.

35

WHAT IS CLAIMED IS:

1.   An enzyme mutant, not found in nature, said
enzyme mutant being derived by the replacement or
modification, in a precursor enzyme, of at least one
catalytic group of an amino acid residue which when in
contact with a selected region of a polypeptide
substrate functions catalytically therewith, to form
an enzyme mutant which is relatively inactive
catalytically with said polypeptide substrate as
compared to said enzyme mutant's catalytic activity
with at least one modified substrate, said modified
substrate being formed by replacing or modifying a
moiety in said selected region to form a modified
moiety which includes said one catalytic group or its
equivalent.

2.   The enzyme mutant of claim 1 wherein said one
catalytic group in said precursor enzyme is
substituted with a second group having a volume which
is less than the volume of said one catalytic group.

3.   The enzyme mutant of claim 1 wherein said
replacement is of a catalytic amino acid residue in
said precursor enzyme with a different amino acid
residue, said catalytic amino acid residue is selected
from the group consisting of His, Lys, Ser, Thr, Cys,
Asp, Glu, Tyr, Met, Phe and Trp and wherein said
different amino acid is selected from the preferred or
alternate amino acid residues of Table I herein.

4.   The enzyme mutant of claim 3 wherein said
modified substrate contains a modified moiety
including a catalytic group or equivalent catalytic
group of Table I herein corresponding to said replaced
catalytic amino acid residue in said precursor enzyme.

-61-

5.   The enzyme mutant of claim 1 wherein said precursor enzyme is selected from the group consisting of oxido-reductases, transferases, hydrolases, lyases, isomerases and ligases.

6.   The enzyme mutant of claim 5 wherein said precursor enzyme is a hydrolase comprising a carbonyl hydrolase.

7.   The enzyme mutant of claim 6 wherein said carbonyl hydrolase is subtilisin.

8.   The enzyme mutant of claim 7 wherein said replaced or modified amino acid residue in said subtilisin is His-64 in B. amyloliquefaciens subtilisin.

9.   The enzyme mutant of claim 8 wherein said His-64 is replaced by Ala.

10. The enzyme mutant of claim 8 wherein said modified substrate contains a modified moiety located at residue P2 of said modified substrate.

11. The enzyme mutant of claim 10 wherein said modified substrate is formed by replacing said moiety at position P2 with histidine.

12. A mutant DNA sequence encoding the enzyme mutant of claim 1.

13. An expression vector containing the DNA of claim 12.

14. Host cells transformed with the expression vector of claim 13.

15. A catalytically active enzyme-substrate complex comprising an enzyme mutant in contact with a modified substrate, wherein said enzyme mutant is not found in nature and is derived by the replacement or

5 modification, in a precursor enzyme, of at least one catalytic group of an amino acid residue which when in contact with a selected region of a substrate functions catalytically therewith, to form said enzyme mutant which is relatively inactive catalytically with

10 said substrate as compared to said enzyme mutant's catalytic activity with at least one said modified substrate, said modified substrate being formed by replacing or modifying a moiety in said selected region to form a modified moiety which includes said

15 one functional group or its equivalent.

16. The enzyme substrate complex of claim 15 wherein said one catalytic group in said precursor enzyme is substituted with a second group having a volume which

20 is less than the volume of said one catalytic group.

17. The enzyme-substrate complex of claim 15 wherein said replacement is of a catalytic amino acid residue in said precursor enzyme with a different amino acid,

25 said catalytic amino acid residue is selected from the group consisting of His, Lys, Ser, Thr, Cys, Asp, Glu, Tyr, Met, Phe, Trp, Asn, Gln and Arg and wherein said different amino acid residue is selected from preferred or alternate amino acid residues of Tables I

30 or II herein.

18. The enzyme-substrate complex of claim 15 wherein said modified substrate contains a modified moiety including a catalytic group or equivalent catalytic

35 group of Tables I or II herein corresponding to said

replaced catalytic amino acid residue in said precursor enzyme.

19. The enzyme-substrate complex of claim 15 wherein said precursor enzyme is selected from the group consisting of oxido-reductases, transferases, hydrolases, lyases, isomerases and ligases.

20. The enzyme-substrate complex of claim 19 wherein said precursor enzyme is a hydrolase comprising a carbonyl hydrolase.

21. The enzyme-substrate complex of claim 20 wherein said carbonyl hydrolase is subtilisin.

22. The enzyme-substrate complex of claim 21 wherein said replaced or modified amino acid residue in said subtilisin is His-64 in B. amyloliquefaciens subtilisin.

23. The enzyme-substrate complex of claim 22 wherein said His-64 is replaced by Ala.

24. The enzyme-substrate complex of claim 22 wherein said modified substrate contains a modified moiety located at residue P2 of said modified substrate.

25. The enzyme-substrate complex of claim 24 wherein said modified substrate is formed by replacing said moiety at position P2 with histidine.

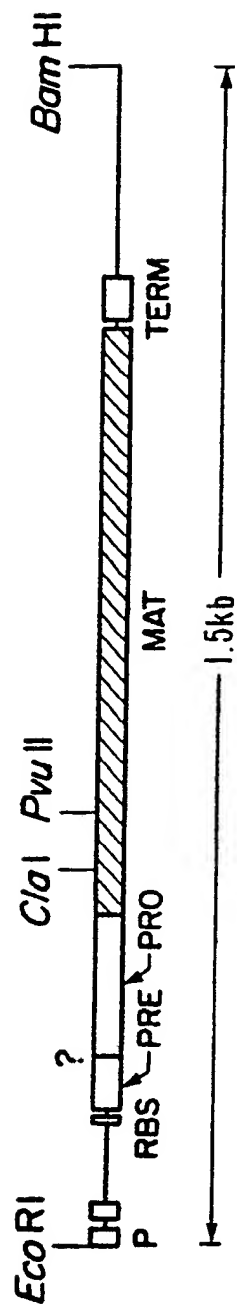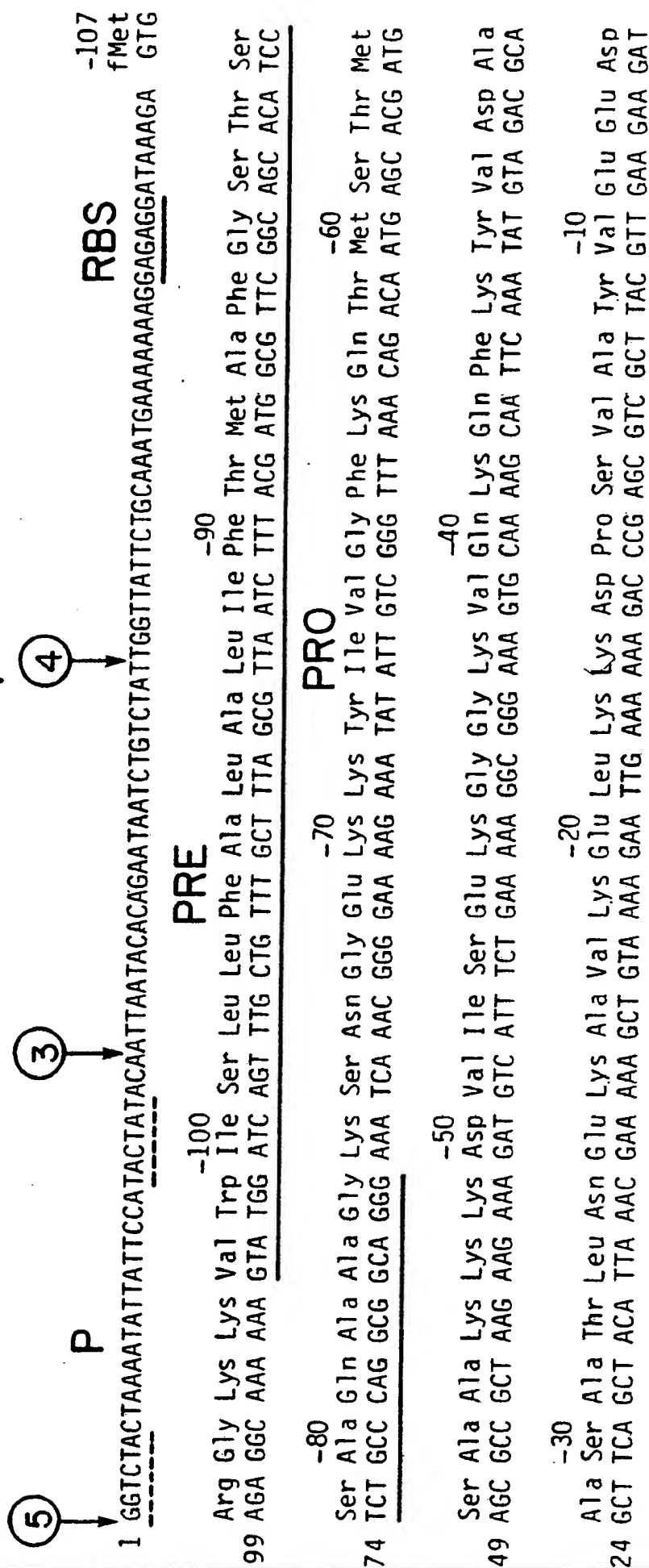26. A process comprising contacting an enzyme mutant and a modified substrate to produce the substrate assisted catalysis of said modified substrate, wherein said enzyme mutant is not found in nature and is derived by the replacement or modification, in a

-64-

precursor enzyme, of at least one catalytic group of
an amino acid residue which when in contact with a
selected region of a substrate functions catalytically
therewith, to form said enzyme mutant which is
5 relatively inactive with said substrate as compared to
said enzyme mutant's catalytic activity with at least
said modified substrate, said modified substrate being
formed by replacing or modifying a moiety in said
selected region to form a modified moiety which
10 includes said one catalytic functional group or its
equivalent.

27. The process of claim 26 wherein said one
catalytic group in said precursor enzyme is
15 substituted with a second group having a volume which
is less than the volume of said one catalytic group.

28. The process of claim 26 wherein said replacement
is of a catalytic amino acid residue in said precursor
20 enzyme with a different amino acid, said catalytic
amino acid residue is selected from the group
consisting of His, Lys, Ser, Thr, Cys, Asp, Glu, Tyr,
Met, Phe, Trp, Asn, Gln and Arg and wherein said
different amino acid residue is selected from
25 preferred or alternate amino acid residues of Tables I
or II herein.

29. The process of claim 28 wherein said modified
substrate contains a modified moiety including a
30 catalytic group or equivalent catalytic group of
Tables I or II herein corresponding to said replaced
catalytic amino acid residue in said precursor enzyme.

30. The process of claim 27 wherein said precursor
35 enzyme is selected from the group consisting of

-65-

oxido-reductases, transferases, hydrolases, lyases, isomerases and ligases.

31. The process of claim 30 wherein said precursor enzyme is a hydrolase comprising a carbonyl hydrolase.

32. The process of claim 31 wherein said carbonyl hydrolase is subtilisin.

33. The process of claim 32 wherein said replaced or modified amino acid residue in said subtilisin is His-64 in B. amyloliquefaciens subtilisin.

34. The process of claim 33 wherein said His-64 is replaced by Ala.

35. The process of claim 33 wherein said modified substrate contains a modified moiety located at residue P2 of said modified substrate.

36. The process of claim 35 wherein said modified substrate is formed by replacing said moiety at position P2 with histidine.

FIG. 1a.

EcoRI ClaI PvuII   BamHI

P RBS PRE PRO MAT TERM

?

1.5kb

FIG. 1b. (top)

P

⑤→
1 GGTCTACTAAAATATTATTCCATACTATACAATTAATACACAGAATAATCTGTCTATTGGTTATTCTGCAAATGAAAAAAGGAGAGGATAAAGA

③→   ④→

RBS

-107
fMet
GTG

PRE

  -100   -90
 Arg Gly Lys Lys Val Trp Ile Ser Leu Leu Phe Ala Leu Ile Phe Thr Met Ala Phe Gly Ser Thr Ser
99 AGA GGC AAA AAG GTA TGG ATC AGT TTG CTG TTT GCT TTA ATC TTT ACG ATG GCG TTC GGC AGC ACA TCC

 -80   -70   -60
 Ser Ala Gln Ala Ala Gly Lys Ser Asn Gly Glu Lys Lys Tyr Ile Val Gly Phe Lys Gln Thr Met Ser Thr Met
74 TCT GCC CAG GCA GCG GCA GGG AAA TCA AAC GGG GAA AAG AAA TAT ATT GTC GGG TTT AAA CAG ACA ATG AGC ACG ATG

PRO

  -50   -40
 Ser Ala Ala Lys Lys Lys Asp Val Ile Ser Glu Lys Gly Gly Lys Val Gln Phe Lys Tyr Val Asp Ala
49 AGC GCC GCT AAG AAG AAA GAT GTC ATT TCT GAA AAA GGG AAA GTG CAA TTC AAA TAT GTA GAC GCA

 -30   -20   -10
 Ala Ser Ala Thr Leu Asn Glu Lys Ala Val Lys Glu Leu Lys Lys Asp Pro Ser Val Ala Tyr Val Glu Glu Asp
24 GCT TCA GCT ACA TTA AAC GAA AAA GCT GTA AAA GAA TTG AAA AAA GAC CCG AGC GTC GCT TAC GTT GAA GAA GAT

```
            -1 |  1 →MAT
                                            10
      His Val Ala His Ala Tyr Ala Gln Ser Val Pro Tyr Gly Val Ser Gln Ile Lys Ala Pro Ala Leu His Ser Gln
399   CAC GTA GCA CAT GCG TAC GCG CAG TCC GTG CCT TAC GGC GTA TCA CAA ATT AAA GCC CCT GCT CTG CAC TCT CAA

                     20              30              40
      Gly Tyr Tyr Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp Ser Gly Ile Asp Ser Ser His Pro Asp Leu Lys Val Ala
474   GGC TAC ACT GGA TCA AAT GTT AAA GTA GCG GTT ATC GAC AGC GGT ATC GAT TCT,CAT CCT GAT TTA AAG GTA GCC

                                    Pro Asn        60 Asp
               50
      Ala Gly Gly Ala Ser Met Val Pro Ser Glu Thr Pro Asn Phe Gln Asp Asn Asn Ser His Gly Thr His Val Ala
549   GCA GGA GGC GCC AGC ATG GTT CCT TCT GAA ACA CCT TTC CAA GAC AAC AAC TCT CAT GGA ACT CAC GTT GCC

                                                   Ser Ala   90
      Gly Thr Val Ala Ala Leu Asn Asn Ser Ile Gly Val Leu Gly Val Ala Pro Ser Ala Ser Leu Tyr Ala Val Lys
624   GGC ACA GTT GCG GCT CTT AAT AAC TCA ATC GGT GTA TTA GGC GTT GCG CCA AGC GCA TCA CTT TAC GCT GTT AAA
          70                          80

                        Asp Ala 100                      110
      Val Leu Gly Ala Asp Gly Ser Gly Gln Tyr Ser Trp Ile Ile Asn Gly Ile Glu Trp Ala Ile Ala Asn Asn Met
699   GTT CTC GGT GCT GAC GGT TCC GGC CAA TAC AGC TGG ATC ATT AAC GGA ATC GAG TGG GCG ATC GCA AAC AAT ATG

                                                   130                          140
      Asp Val Ile Asn Met Ser Leu Gly Gly Pro Ser Gly Ser Ala Ala Leu Lys Ala Ala Val Asp Lys Ala Val Ala
774   GAC GTT ATT AAC ATG TCT CTC GGT GGA CCT TCT GGT TCT GCT GCG CTT AAA GCA GTT GAT AAA GCC GTT GCA
          120

                                                   Ser Thr 160
      Ser Gly Val Val Val Val Ala Ala Ala Gly Asn Glu Gly Thr Ser Gly Ser Ser Ser Thr Val Gly Tyr Pro Gly
849   TCC GGC GTA GTT GTT GCG GCA GGT AAC GAA GGT ACT TCC GGC AGC AGC ACA GTG GGC TAC CCT GGT
          150
```

FIG. 1b.(mid.)

```
            170                         180                           190
      Lys Tyr Pro Ser Val Ile Ala Val Gly Ala Val Asp Ser Ser Asn Gln Arg Ala Ser Phe Ser Val Gly Pro
924   AAA TAC CCT TCT GTC ATT GCA GTA GGC GCT GTT GAC AGC AGC AAC CAA AGA GCA TCT TTC TCA AGC GTA GGA CCT

                                    200                                210
      Glu Leu Asp Val Met Ala Pro Gly Val Ser Ile Gln Ser Thr Leu Pro Gly Asn Lys Tyr Gly Ala Tyr Asn Gly
999   GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC GGT

            220                               230                              240
      Thr Ser Met Ala Ser Pro His Val Ala Gly Ala Ala Ala Leu Ile Leu Ser Lys His Pro Thr Trp Thr Asn Thr
074   ACG TCA GCA TCT CCG CAC GTT GCC GGA GCG GCT TTG ATT CTT TCT AAG CAC CCG ACA AAC AAC ACT

                                   250   Gln                            260
      Gln Val Arg Ser Ser Leu Glu Asn Thr Thr Thr Lys Leu Gly Asp Ser Phe Tyr Tyr Gly Lys Gly Leu Ile Asn
149   CAA GTC CGC AGC AGT TTA GAA AAC ACC ACT ACA AAA CTT GGT GAT TCT TTC TAC TAT GGA AAA GGG CTG ATC AAC

            270               275
      Val Gln Ala Ala Ala Gln OC
224   GTA CAG CAG GCG GCA GCT CAG TAA
```

TERM

```
      AACATAAAAAACCGGCCTTGGCCCCGCGGGTTTTTTATTATTTTTCTTCCGCATGTTCAATCCGCTCC

316   ATAATCGACGGATGGCTCCCTCTGAAAATTTTAACGAGAAACGGGGTTGACCCGGCTCAGTCCCGTAACGGCCAAGTCCTGAAACGTCTCAATCGCCG

316   CTTCCCGGTTTCCGGTCAGCTCAATGCCGTAACGGTCGGCGGGCGTTTCCTGATACCGGGAGACGGCATTCGTAATCGGATC
```

FIG. 1b. (bottom)
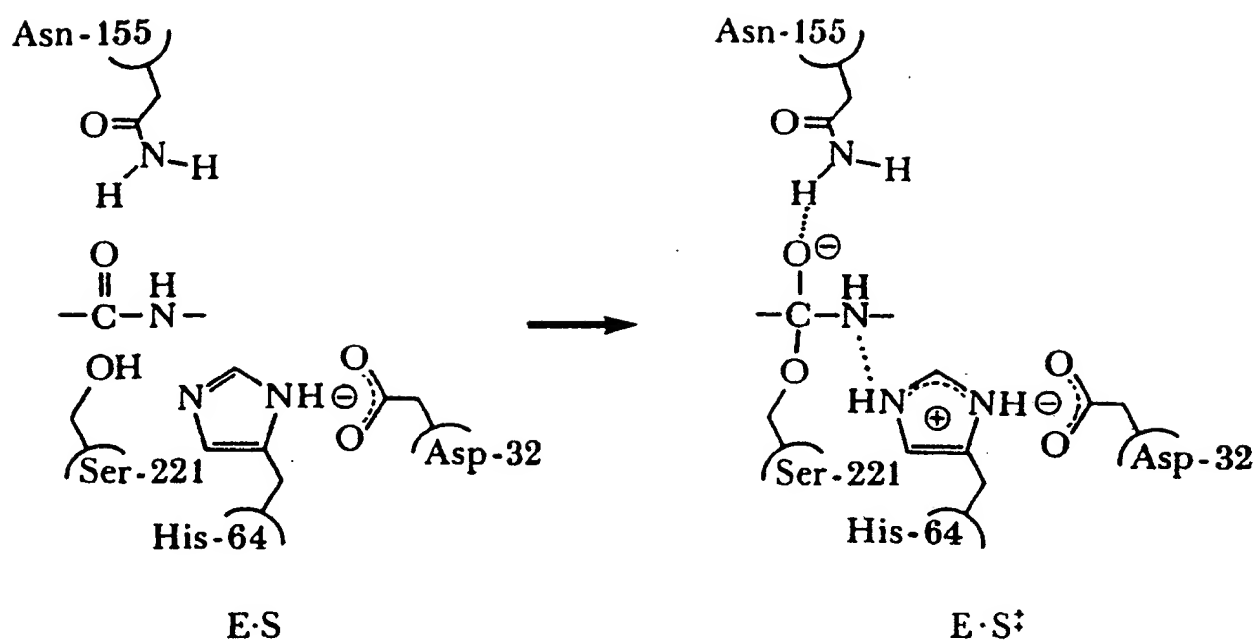
FIG. 2.

pH Profile for Cys24Ala64 Against sucFAHFpna
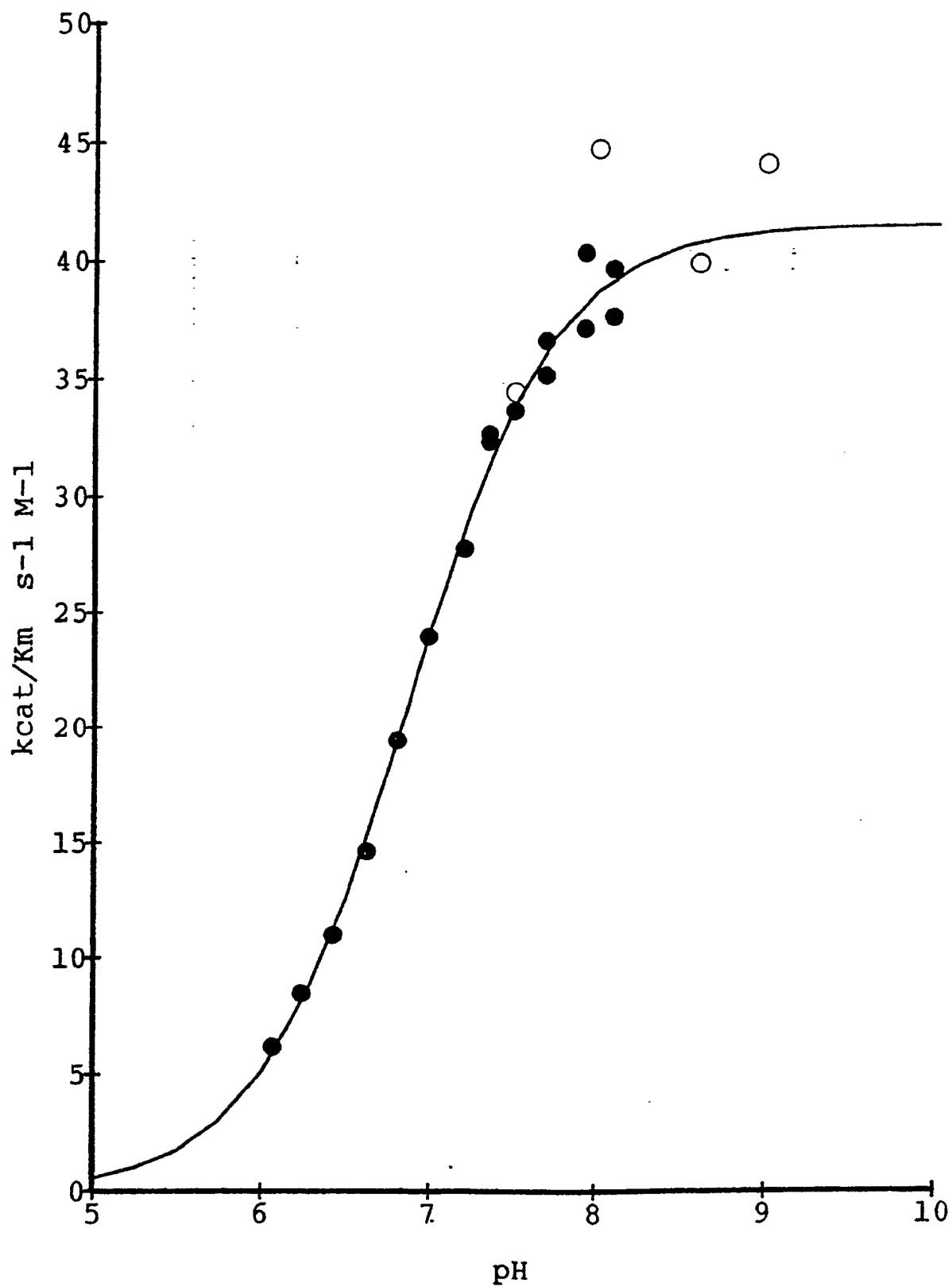
FIG. 2a.

pH Profile for Cys24Ala64 against sucFAAFpna
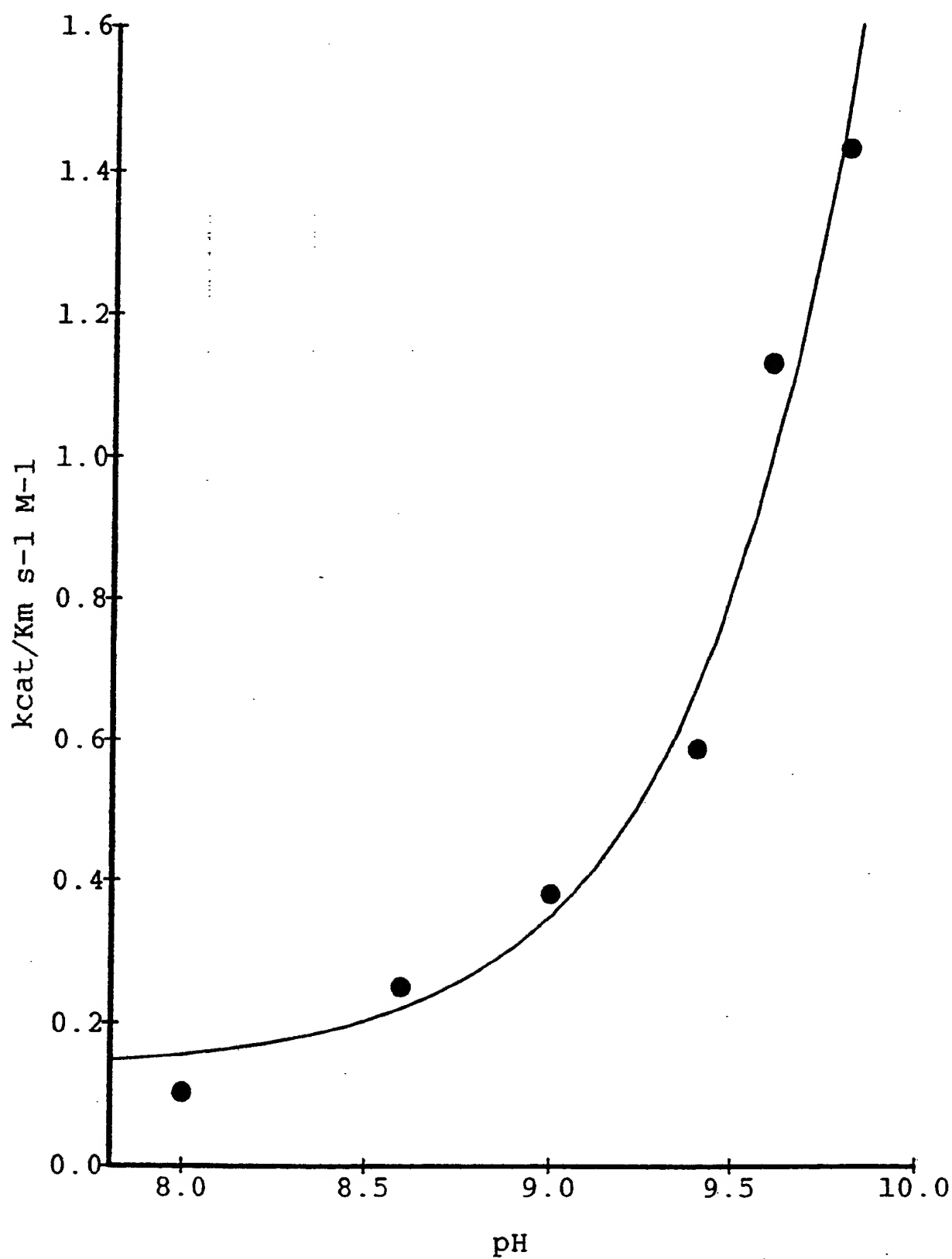
FIG. 2b.

Hydroxide Ion Dependence of C24A64 Against SucFAAFpna
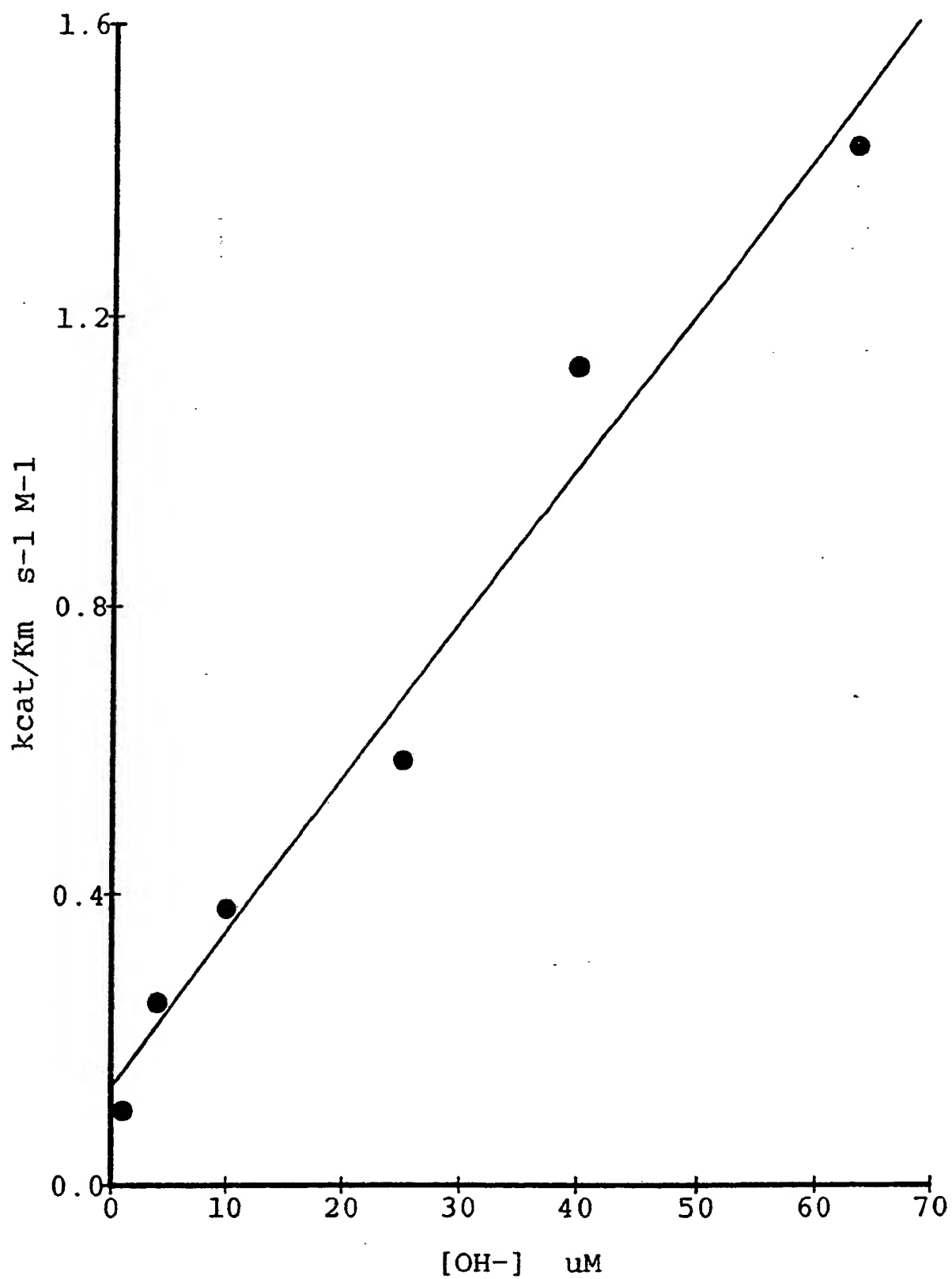
FIG. 2c.

Cys24Ala64

Cys24
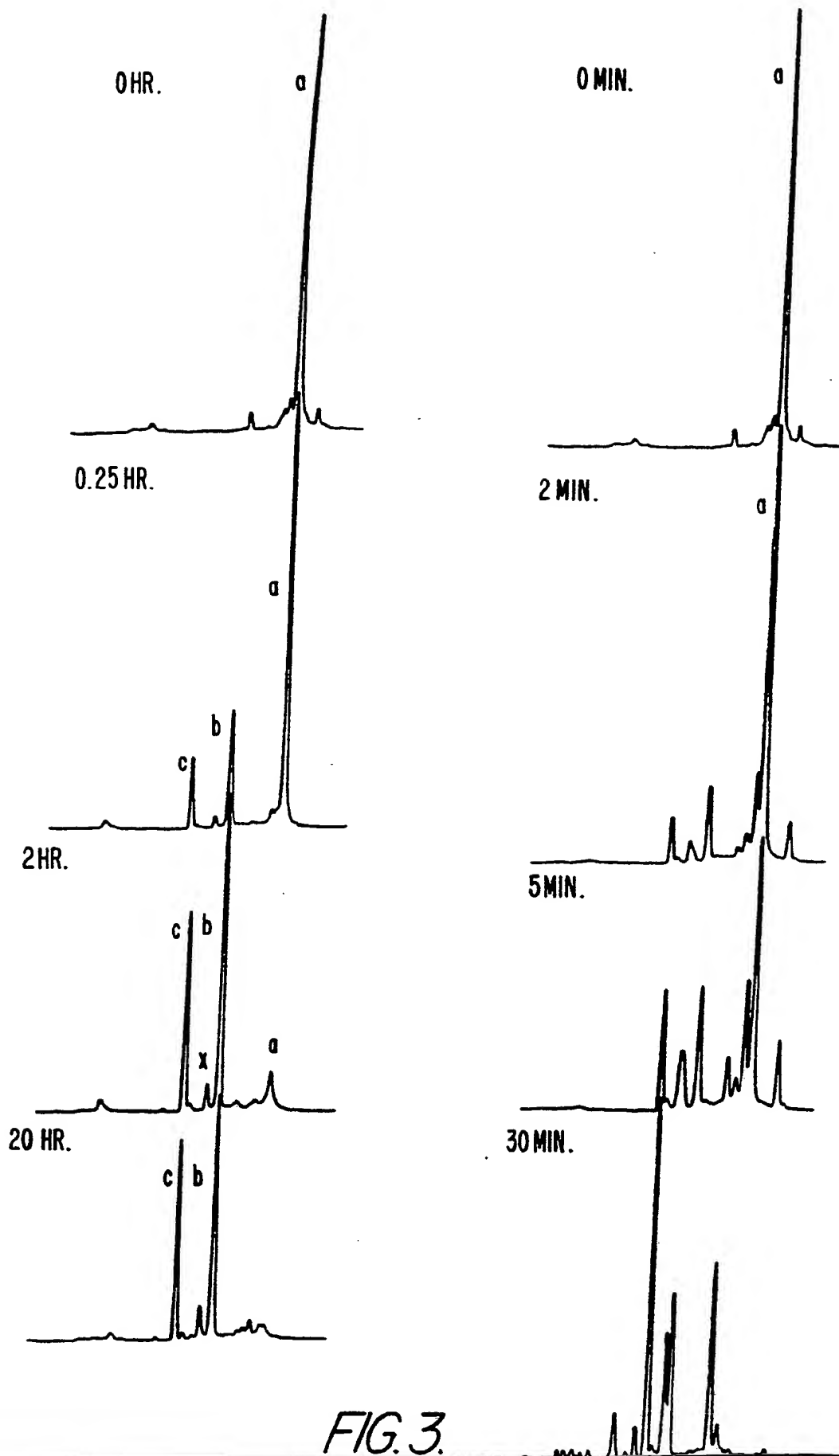
FIG. 3.

FIG. 3a.

Homology of Bacillus proteases
1. Bacillus amyloliquifaciens
2. Bacillus subtilis var. I168
3. Bacillus licheniformis (carlsbergensis)

```
              1                    10                  20
1   A  Q  S  V  P  Y  G  V  S  Q  I  K  A  P  A  L  H  S  Q  G
2   A  Q  S  V  P  Y  G  I  S  Q  I  K  A  P  A  L  H  S  Q  G
3   A  Q  T  V  P  Y  G  I  P  L  I  K  A  D  K  V  Q  A  Q  G

             21                   30                  40
1   Y  T  G  S  N  V  K  V  A  V  I  D  S  G  I  D  S  S  H  P
2   Y  T  G  S  N  V  K  V  A  V  I  D  S  G  I  D  S  S  H  P
3   F  K  G  A  N  V  K  V  A  V  L  D  T  G  I  Q  A  S  H  P

             41                   50                  60
1   D  L  K  V  A  G  G  A  S  M  V  P  S  E  T  N  P  F  Q  D
2   D  L  N  V  V  G  G  A  S  F  V  P  S  E  T  N  P  Y  Q  D
3   D  L  N  V  A  G  G  A  S  F  V  A  G  E  A  Y  N  T  *  D
```

```
                 70                    80
61  N  S  H  T  G  H  V  T  G  A  V  A  L  N  N  S  I  G
    S  S  H  T  G  H  V  T  G  A  I  A  L  N  N  S  I  G
    N  G  H  T  G  H  V  T  G  A  V  A  L  D  N  T  T  G

                 90                    100
81  V  L  G  V  A  P  S  A  S  L  Y  A  V  K  V  L  G  A  D  G
    V  L  G  V  S  P  S  A  S  L  Y  A  V  K  V  L  D  S  T  G
    V  G  G  V  A  P  S  V  S  L  Y  A  V  K  V  L  N  S  S  G

                 110                   120
101 S  G  Q  Y  S  W  I  I  N  G  I  E  W  A  I  A  N  N  M  D
    S  G  Q  Y  S  W  I  I  N  G  I  E  W  A  I  S  N  N  M  D
    S  S  S  Y  S  G  I  V  S  G  I  E  W  A  T  T  N  G  M  D
.PA

                 130                   140
121 V  I  N  M  S  L  G  G  S  P  G  S  A  A  L  K  A  A  V  D
    V  I  N  M  S  L  G  G  T  P  G  S  T  T  L  K  T  V  V  D
    V  I  N  M  S  L  G  G  S  A  G  S  A  A  M  M  Q  A  V  D
```

FIG. 3a. (cont.)

Protein sequence alignment (three aligned sequences), positions 141–220:

```
141  K  A  V  A  S  G  V  V  V  A  A  G  N  E  G  T  S  G      160
     K  A  V  S  S  G  I  V  A  A  A  G  N  E  G  S  S  G
     N  A  Y  A  R  G  V  V  V  A  A  G  N  S  G  N  S  G

161  S  S  T  V  T  G  V  K  Y  G  P  A  V  I  A  A  V  A  V   180
     S  T  S  V  T  G  V  K  Y  A  P  A  V  I  A  A  G  V  V
     S  N  T  I  N  I  G  K  Y  D  P  V  I  A  A  G  V  V

181  D  N  Q  R  A  S  F  S  A  S  V  G  P  E  L  D  V  M  A   200
     N  N  Q  R  A  S  F  S  A  A  G  P  E  L  D  V  M  A
     D  S  N  R  A  S  F  S  V  G  A  E  L  E  A  V  M  A

201  P  V  S  I  Q  S  T  L  P  G  K  N  K  Y  G  A  N  G  T   220
     P  V  S  I  Q  S  T  L  P  G  G  Y  G  Y  A  Y  N  G  T
     P  A  G  V  Y  S  T  Y  P  T  N  A  T  Y  Y  L  A  N  G  T
```

FIG. 3a. (cont.)

FIG. 3a.(cont.)

```
221        230             240
S M A S P H V A A G A A L I L S K H P N
S M A T P H V A A G A A L H L S K H P T
S M A S P H V A A G A A L I L S K H P N

241        250             260
W N T Q V R S R S L E N S T T K L G D S
W N A Q V R D R R L E S T A T Y L G N S
L A S Q V R N R R L S S S S T Y L G S S

261        270
F Y G K G L I N V V Q A A Q
F Y G K G L H N V Q Q A A Q
F Y G K G L I N V E E A A Q
```

ALIGNMENT OF B.AMYLOLIQUIFACIENS SUBTILISIN AND THERMITASE
1. B.amyloliquifaciens subtilisin
2. thermitase

FIG. 3b.

```
             1                          10                         20
1  A  Q  S  V  *  P  Y  *  *  *  *  G  V  S  Q  I  K  A  P  A  L  H  S  Q  G  Y  T  G  S
2  Y  T  P  N  D  P  Y  *  S  Q  R  Q  G  P  I  Q  A  Q  A  W  D  I  A  E  G  S  G  Y  T

             30                         40                         50
1  N  V  K  V  A  V  I  D  S  G  I  D  S  S  H  P  D  L  *  K  V  A  G  G  A  S  M  V  P
2  G  K  G  V  K  V  A  V  L  D  T  G  I  S  T  *  H  P  D  L  N  I  R  G  G  A  S  F  V

             60                         70                         80                 90
1  *  *  *  S  E  T  N  P  F  Q  D  N  N  S  H  G  T  H  V  A  G  T  V  A  A  L  N  N  S  I  G  V  L  G  V  A  P  S  A  S  L
2  P  G  E  P  S  T  Q  *  Q  D  G  N  *  G  N  G  H  G  T  H  V  A  G  T  I  A  A  L  N  N  S  I  G  V  L  G  V  A  P  N  A  E  L
```

FIG. 3b. (cont.)

```
                    200                          210
S  V  G  P  E  L  D  V  M  A  P  G  V  S  I  Q  S  T  L  P
T  Y  G  S  V  V  D  V  A  A  P  G  S  W  I  Y  S  T  Y  P

                    220                          230
G  N  K  Y  G  A  Y  N  G  T  S  A  P  H  V  A  G  A
T  S  T  S  L  S  L  S  T  S  M  A  T  P  V  A  G  V

                    240                          250
A  L  L  K  S  Q  V  T  N  W  P  H  V  R  S  L
A  L  S  Q  R  G  S  A  S  *  R  I  A  A  I

                              260
E  T  T  K  *  L  G  D  S  F  Y  G  K  G  L  N
E  T  D  K  I  S  G  T  G  Y  Y  A  G  R  V  N

270
V  Q  A  A  Q
A  Y  K  A  V           Y  Q
```

FIG. 3b. (cont.)

FIG. 3c.

TOTALLY CONSERVED RESIDUES IN SUBTILISINS

141
S G   P A   G
150   G G   Y P   N

161
170   Y P   V

181
190   S S   F S
200   V

201
P G   S F   G
210

220
G T

221
M A   P H   V A   G
230

240
241   R

250

261
270   N

FIG.3c.(cont.)

FIG. 4.

FIG. 5.

B. AMYLOLIQUEFACIENS SUBTILISN ————

PEPTIDE SUBSTRATE —·—·—·—

CATALITIC TRIAD (Asp^{32}, His^{64}, Ser^{221}) —·—·—·—

His P2

Asp^{32}

His^{64}

Tyr P1

Ser^{221}

SUBSTRATE—SUBTILISIN

FIG. 6.

FIG.7a.

FIG.7c.

FIG.7b.

Cys24Ala64
0HR.   a

0.25 HR.   a
b
c

2 HR.   b
c
x

20 HR.   b
x

FIG. 8a.

Cys24
0MIN.   a

2 MIN.   a

5 MIN.

30 MIN.

FIG. 8b.

BOVINE TRYPSIN —————

BOVINE PANCREATIC TRYPSIN INHIBITOR (BPTI) — — —

CATALYTIC TRIAD OF TRYPSIN $(Ser^{195}, His^{57}, Asp^{102})$ — · — · —

PTI — TRYPSIN
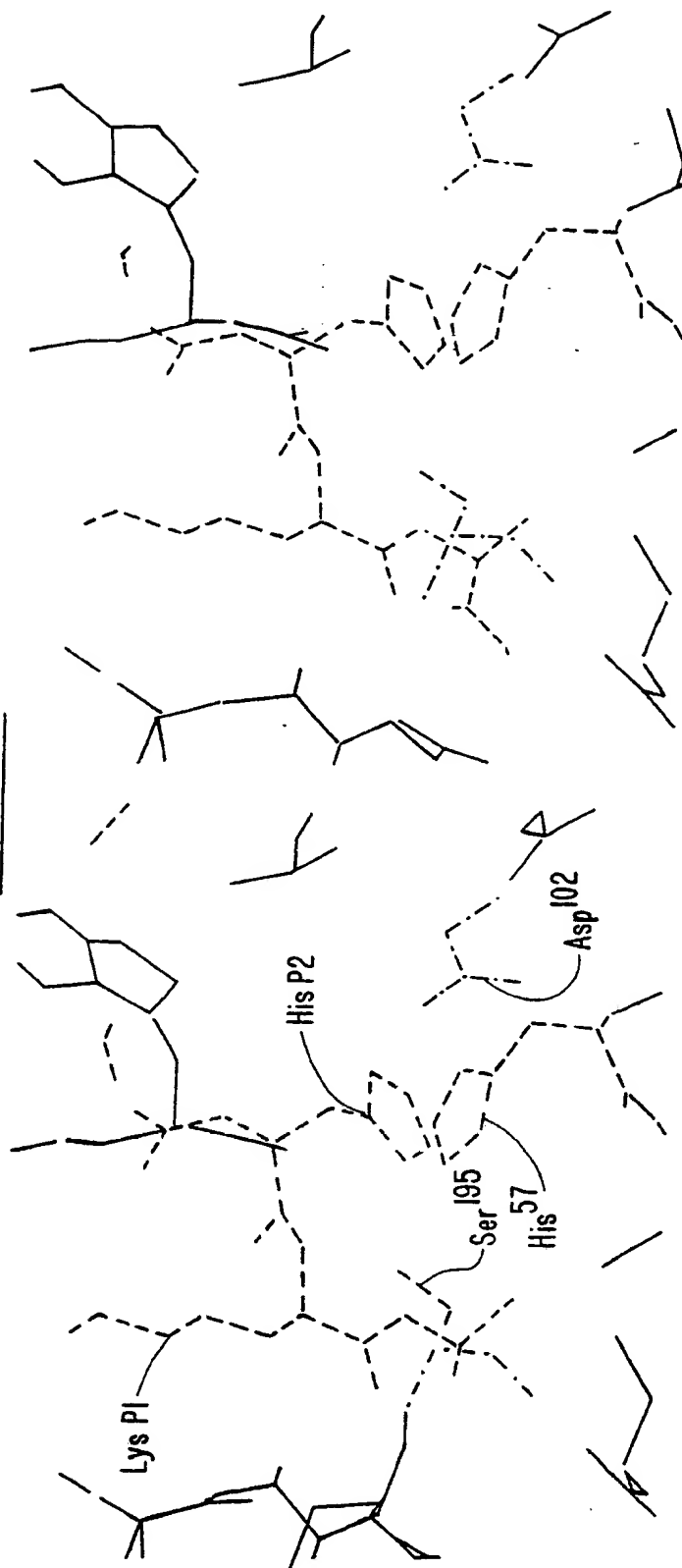
His P2

Asp^{102}

Ser^{195}

His^{57}

Lys PI

FIG. 9.

INTERNATIONAL SEARCH REPORT

I. CLASSIFICATION OF SUBJECT MATTER (If several classification symbols apply, indicate all) *

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[4]:C 12 N 15/00; C 12 N 9/00; C 12 N 9/54; C 12 N 9/96

II. FIELDS SEARCHED

Minimum Documentation Searched [7]	
Classification System	Classification Symbols
IPC[4]	C 12 N; C 12 P

Documentation Searched other than Minimum Documentation
to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with Indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
P,X	Science, volume 237, 24 July 1987, (Washington, DC, US), P. Carter et al.: "Engineering enzyme specificity by "Substrate-Assisted Catalysis" ", pages 394-399	1-36
A	Proc. Natl. Acad. Sci. USA, volume 84, March 1987, (Washington, DC, US), J.A. Wells et al.: "Designing substrate specificity by protein engineering of electrostatic inter-actions", pages 1219-1223	

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
17th June 1988	1 2 JUIL. 1988

International Searching Authority	Signature of Authorized Officer
EUROPEAN PATENT OFFICE	P.C.G. VAN DER PUTTEN

Form PCT/ISA/210 (second sheet) (January 1985)